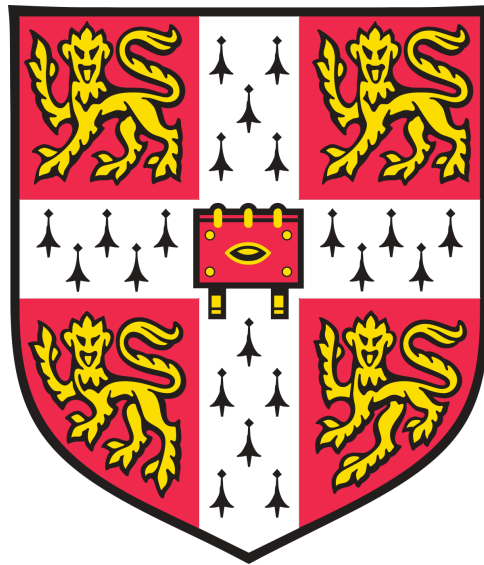


The Cell-Essentiality of KAT7 in Acute Myeloid Leukemia



Yan Zi Au

**University of Cambridge
Goville and Caius College**

*This dissertation is submitted for the degree of
Doctor of Philosophy
August 2019*

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. No substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution. It does not exceed the prescribed word limit of 60,000 words.

Yan Zi Au
August 2019

Abstract

Acute myeloid leukemia (AML) is an aggressive hematopoietic malignancy, characterized by the uncontrolled proliferation and differentiation arrest of myeloid progenitors. Chemotherapy has been the front-line treatment for decades and cure remains elusive for the majority of AML patients. Genome-wide CRISPR-Cas9 screens have previously identified *KAT7* as an AML-specific cell-essential gene and therefore may represent a potential novel therapeutic target for AML. Here, I show that *KAT7* loss leads to a rapid and dramatic global reduction in both H3K14ac and H4K12ac in association with reduced proliferation, increased apoptosis or enhanced differentiation of AML cells driven by the translocation of *Mixed-lineage leukemia (MLL)* gene. Mice transplantation with *KAT7* knock-out AML cell line showed delayed disease progression and prolonged survival compared to those injected with the wild-type counterpart. The acetyltransferase activity of *KAT7* is essential for MLL-fusion AML as the E508Q catalytic dead mutant is unable to sustain the leukemic programme. Using the auxin-inducible degron (AID) system to induce rapid *KAT7* protein degradation, I showed that *KAT7* is required for the recruitment of the MLL-fusion associated adaptor proteins such as BRD4 and AF4 to gene promoters, which are critical for the maintenance of the MLL-AF9 transcriptional programme. Although not found to be mutated among cases of AML, *KAT7* is a plausible therapeutic target for this poor prognosis subtype of AML.

Acknowledgement

PhD is not a solo journey. I am blessed with the presence of many wonderful humans that made the toughest days of PhD bearable. First and foremost, I would like to thank the Wellcome Trust for giving me the opportunity to pursue my passion in translational research at the University of Cambridge. I am thankful to Gonville and Caius and the weekly MDRs which allowed me to meet some of the greatest minds.

My gratitude goes to my advisor, Brian Huntly, who has always made time for discussion since the start of my PhD— a mentor like no other. I am thankful to George Vassiliou for his undiminishing encouragement and guidance throughout my PhD and to Kosuke Yusa for teaching me the need to stand up for myself through very unconventional ways.

Endless hours in the lab was accompanied by supportive, encouraging and knowledgeable teammates, whom I could not have survived without. Special shout out to Muxin Gu and Swee Hoe Ong who has helped me immensely in the bioinformatics analyses. Millions of thanks to Gonja Gozdecka, who has blessed me with her wisdom on ChIP experiments and her positive outlook for every situation. Jason Yu, my role model in science, has shown me how to be a critical but optimistic scientist. Kasia Tilgner and Shanade Dunn deserves special mention for being there for me in the good and the not so good times. My thanks also go to Etienne De Braekeleer who has helped with the in vivo experiments and to Bee Ling Ng, Jennie Graham, Chris Hall and Sam Thompson for their support in flow cytometry. I am thankful to Josep Nomdedeu for his inspirational words and vision. The Sanger-Cambridge experience wouldn't be the same without, Yan Shao, Gene Koh, Nicholas Lee, Sebastian Grossman, Eddie Cano Gamez, Josh Neoh and Chin Mei Lee. The kind words of these wonderful friends and colleagues have lifted me from some of the darkest days.

I would not be where I am today without people whom I've met in the past. I am forever indebted to Max Reuter, Babis Rallis and Jugh Bahler for their mentorship and the research opportunities when I was still an undergrad. To friends who have believed in me for a decade: JJ Leow, Jeremy Tan, Rachel Tay, Jia Yi Gan, May Yee and Rahul Sakhrani- I am glad that we can grow old without growing apart.

My deep appreciation goes to my better half, Edward Kah Wei Tan, whom I have the great fortune to meet and grow with during my time in Cambridge. He has nurtured me with kindness, compassion and patience. Our journey together continues in the other Cambridge!

There is truly no place like home. I am eternally grateful and appreciative to my family, grandparents, aunts, uncles and cousins who have constantly supported and believed in me in every step. My lovely parents, brother Zher Wen and sister Yan Yi are my pillar of strength in all and every aspects of life since the day I was born. Their nutritious love for me is boundless and shapes me into the person I am today. Words are inadequate to express my love for them and because of them my heart is wholesome!

Table of Contents

| | |
|--|-----------|
| Abstract | 4 |
| Acknowledgement | 5 |
| List of Figures..... | 11 |
| List of Tables..... | 13 |
| List of Publication..... | 13 |
| List of Abbreviations | 14 |
| Chapter 1: Introduction | 17 |
| 1.1 Acute myeloid leukaemia and disease pathogenesis..... | 17 |
| 1.2 Current and emerging treatments for AML..... | 18 |
| 1.3 CRISPR-Cas9 as a powerful tool to identify vulnerabilities in AML | 19 |
| 1.4 KAT7 protein complexes and histone tail specificity | 21 |
| 1.5 Molecular, cellular and biological functions of KAT7 | 22 |
| 1.5.1 DNA-replication | 22 |
| 1.5.2 Transcription and regulation of gene expression | 23 |
| 1.5.3 Physiological roles of KAT7 | 23 |
| 1.6 KAT7 in cancer and diseases | 24 |
| 1.7 Aims | 24 |
| Chapter 2: Materials and methods | 25 |
| 2.1 Cell culture..... | 25 |
| 2.2 gRNA design and cloning of gRNA into expression vector..... | 26 |
| 2.3 Virus production and transduction | 26 |
| 2.4 Generating Cas9 expressing Nomo-1 cell line and blasticidin selection..... | 26 |
| 2.5 Cell sorting and FACS analyses..... | 28 |
| 2.6 Sub-cloning and characterisation of Cas9-expressing AML cells..... | 28 |
| 2.7 Functional assays | 28 |
| 2.7.1 Proliferation | 28 |
| 2.7.2 Differentiation..... | 28 |
| 2.7.3 Apoptosis..... | 29 |

| | |
|---|-----------|
| 2.7.4 Cell cycle..... | 29 |
| 2.8 Cloning and generation of gRNA-resistant wild-type, G485A and E508Q mutant KAT7 cell lines..... | 29 |
| 2.9 Sanger Sequencing | 31 |
| 2.10 Western blot..... | 31 |
| 2.11 In vivo mouse work..... | 32 |
| 2.12 Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) | 32 |
| 2.13 Optimization of qPCR primers..... | 33 |
| 2.14 Optimization for ChIP-seq | 33 |
| 2.15 Validation of antibodies used in ChIP-seq..... | 37 |
| 2.16 ChIP-Seq sample preparation | 38 |
| 2.17 ChIP-qPCR..... | 39 |
| 2.18 ChIP-seq processing..... | 40 |
| 2.19 RNA extraction and RNA-seq processing..... | 40 |
| 2.20 Data availability | 40 |
| 2.21 Generation of Auxin Inducible Degron (AID) KAT7 protein degradation system and treatment with indole-3-acetic acid (IAA)..... | 41 |
| 2.23 IBET -151 treatment and viability assay | 41 |
| 2.24 Statistical analysis..... | 41 |
| Chapter 3: Phenotypic validation of CRISPR-Cas9 AML-specific cell essential genes identify KAT7 as a potential novel therapeutic target for MLL-fusion AML | 45 |
| 3.1 Introduction | 45 |
| 3.2 Assessing the Cas9 activity of human AML cell lines..... | 45 |
| 3.3 Validation of AML-specific druggable candidate genes from CRISPR-Cas9 dropout screens.... | 49 |
| 3.4 Proliferation of KAT7 and SIK3 knock-out in various AML and CML cell lines..... | 52 |
| 3.5 KAT7 and SIK3 knock-out in MOLM-13..... | 55 |
| 3.5.1 Apoptosis..... | 55 |
| 3.5.2 Differentiation..... | 56 |
| 3.5.3 Cell cycle..... | 59 |

| | |
|---|-----------|
| 3.6 <i>KAT7</i> knock-out in additional AML cell lines..... | 60 |
| 3.6.1 Apoptosis..... | 60 |
| 3.6.2 Proliferation | 62 |
| 3.6.3 Differentiation..... | 64 |
| 3.7 In vivo effects of <i>KAT7</i> loss..... | 65 |
| 3.8 Expression of <i>KAT7</i> in normal haematopoiesis | 66 |
| 3.9 <i>KAT7</i> mutations and expression in other cancer types | 68 |
| 3.10 Discussion..... | 70 |
| Chapter 4: Molecular characterization of <i>KAT7</i> in AML..... | 73 |
| 4.1 Introduction | 73 |
| 4.2 <i>KAT7</i> protein and H3K14ac levels are depleted after g <i>KAT7</i> mediated knock-out | 73 |
| 4.3 <i>KAT7</i> modified histone marks H3K14ac and H4K12ac are depleted in <i>KAT7</i> knock-out cells | 74 |
| 4.4 Changes in other histone modifications upon <i>KAT7</i> loss..... | 76 |
| 4.5 <i>KAT7</i> catalytic activity | 77 |
| 4.5.1 G485A catalytic active mutant | 78 |
| 4.5.2 E508Q catalytic dead mutant | 82 |
| 4.5.3 Dominant negative effect of E508Q <i>KAT7</i> on histone acetylation and cell proliferation..... | 87 |
| 4.6 <i>KAT7</i> predominantly occupies the gene promoter region | 91 |
| 4.7 Transcriptomics changes following <i>KAT7</i> loss | 94 |
| 4.8 Discussion | 95 |
| Chapter 5: Mechanistic link between <i>KAT7</i> and the pathogenesis of <i>MLL</i>-fusion AML | 97 |
| 5.1 Introduction | 97 |
| 5.2 Majority of the <i>MLL</i> -AF9 and <i>MLL</i> -AF4 spreading targets are bound by <i>KAT7</i> | 98 |
| 5.3 Top <i>KAT7</i> bound genes are <i>MLL</i> -fusion targets in MOLM-13 and MV4-11 but not in OCI-AML3..... | 100 |
| 5.4 Genes co-bound by <i>MLL</i> -AF9 and <i>KAT7</i> at promoter are downregulated after <i>KAT7</i> loss in MOLM-13 | 101 |
| 5.5 Loss of <i>KAT7</i> results in general gene expression changes in MV4-11 that is not associated with <i>MLL</i> -AF4 spreading targets..... | 104 |
| 5.6 Auxin-inducible degron (AID) system for rapid <i>KAT7</i> protein degradation | 106 |

| | |
|---|------------|
| 5.7 Gene expression changes and the differentiation of IAA treated MOLM-13 AID and MV4-11 AID cells..... | 108 |
| 5.8 Loss of KAT7 and the occupancy of MLL-fusion proteins | 110 |
| 5.8.1 MLL-AF9..... | 111 |
| 5.8.2 MLL-AF4..... | 111 |
| 5.9 Occupancy of BRD4 at MLL-fusion target loci is dependent on KAT7 in MOLM-13 but not in MV4-11 | 114 |
| 5.10 Loci-specific changes in KAT7-mediated histone lysine acetylation levels upon IAA treatment | 116 |
| 5.11 Working model for KAT7 mechanism of action in regulating the expression of MLL-AF9 target genes..... | 119 |
| 5.12 MOLM-13 E508Q sensitizes cells to IBET-151 treatment..... | 120 |
| 5.13 Discussion..... | 121 |
| Chapter 6: Discussion and future work..... | 123 |
| References | 127 |

List of Figures

| | |
|---|----|
| Figure 1 Protein complexes of KAT7 | 21 |
| Figure 2.1 Cas9, Cas9 reporter and gRNA constructs | 27 |
| Figure 2.2 Double digestion of plasmid backbone for the integration of E508Q mutant sequence | 31 |
| Figure 2.3 DNA fragments after sonication of cross-linked MOLM-13..... | 35 |
| Figure 2.4 DNA fragments after sonication of cross-linked MV4-11 and OCI-AML3..... | 36 |
| Figure 2.5 Analysis of input DNA fragments after the sonication of cross-linked MOLM-13 and OCI-AML3 by bioanalyzer | 37 |
| Figure 2.6 Immunoblot of ChIP using KAT7, H3K14ac and H4K12ac antibodies in MOLM-13 | 38 |
| Figure 3.1 Cas9 activity of AML and CML cell lines before single-cell cloning..... | 48 |
| Figure 3.2 Cas9 activity of AML and CML cell lines after single-cell cloning..... | 48 |
| Figure 3.3 Cas9 activity and effect on functional phenotype | 49 |
| Figure 3.4 Validation of candidate hits from the CRISPR-Cas9 dropout screen in MOLM-13 and OCI-AML3..... | 51 |
| Figure 3.5 Proliferation of SIK3 and KAT7 knock-out in additional cell lines | 53 |
| Figure 3.6 Proliferation of MOLM-13 and MV4-11 with different gRNAs targeting KAT7 (gKAT7) and SIK3 (gSIK3)..... | 54 |
| Figure 3.7 Apoptosis of SIK3 and KAT7 knock-out in AML cell lines..... | 55 |
| Figure 3.8 Time course CD11b differentiation assay in SIK3, KAT7 and KAT2A knock-out MOLM-13 cells..... | 57 |
| Figure 3.9 Time course CD13 differentiation assay in SIK3, KAT7 and KAT2A knock-out MOLM-13 cells..... | 58 |
| Figure 3.10 CD11b and CD13 cell surface marker expression of SIK3 and KAT7 knock-out in MOLM-13 | 59 |
| Figure 3.11 Cell cycle analysis of SIK3 and KAT7 knock-out in MOLM-13 | 60 |
| Figure 3.12 Apoptosis of KAT7 knock-out in AML cell lines..... | 61 |
| Figure 3.13 Proliferation of KAT7 knock-out in AML cell lines..... | 62 |
| Figure 3.14 Simulation of cell proliferation with different percentage of initial transduced cells | 63 |
| Figure 3.15 CD11b expression of KAT7 knock-out in AML cell lines..... | 64 |
| Figure 3.16 Expression of myeloid differentiation marker in KAT7 knock-out MV4-11 and THP-1 .. | 65 |
| Figure 3.17 Disease progression and survival of mice injected with luciferase-expressing MOLM-13 cells..... | 66 |

| | |
|---|-----|
| Figure 3.18 Expression of KAT7 in normal hematopoietic cells | 67 |
| Figure 3.19 KAT7 expression in 37 cancer types from The Cancer Genome Atlas (TCGA)..... | 69 |
| Figure 3.20 Location and type of mutations in the protein-coding region of KAT7 found across The Cancer Genome Atlas (TCGA) dataset..... | 70 |
| | |
| Figure 4.1 Changes in levels of KAT7 protein and KAT7-mediated histone lysine acetylation following gRNA-mediated knock-out of KAT7 | 75 |
| Figure 4.2 Levels of activating and inactivating histone marks after KAT7 knock-out | 77 |
| Figure 4.3 Schematic of KAT7 protein and the targeting sites of the different gKAT7 | 78 |
| Figure 4.4 Genetic construct of G485A mutant KAT7 transgene..... | 80 |
| Figure 4.5 G485A mutant possesses catalytic activity and does not affect proliferation and differentiation of MOLM-13 | 81 |
| Figure 4.6 Potential human KAT7 catalytic mutants based on studies in yeast Esa1 | 83 |
| Figure 4.7 Genetic construct of E508Q mutant KAT7 transgene. | 84 |
| Figure 4.8 E508Q catalytic dead mutant KAT7 affects the proliferation and differentiation of MOLM-13 | 87 |
| Figure 4.9 CRISPR-Cas9 depletion and gene expression of the members within the KAT7 protein complex..... | 88 |
| Figure 4.10 E508Q KAT7 mutant MOLM-13 populations with different levels of GFP expression..... | 90 |
| Figure 4.11 Genome-wide chromatin binding of KAT7 in MOLM-13, OCI-AML3 and MV4-11 | 93 |
| Figure 4.12 Gene expression changes after KAT7 knock-out | 95 |
| | |
| Figure 5.1 Shared spreading targets of MLL-AF9 and MLL-AF4 | 99 |
| Figure 5.2 Majority of MLL-AF9 and MLL-AF4 spreading targets are co-bound by KAT7 at the promoter | 100 |
| Figure 5.3 Ranked normalized KAT7 promoter occupancy signal..... | 101 |
| Figure 5.4 Transcriptomics change after KAT7 knock-out and the chromatin occupancy of KAT7 in MOLM-13 and OCI-AML3..... | 103 |
| Figure 5.5 Transcriptomics changes in MV4-11 after KAT7 knock-out and the chromatin occupancy of KAT7..... | 105 |
| Figure 5.6 Auxin-inducible degron (AID) MOLM-13 and MV4-11 induces rapid KAT7 protein degradation and H3K14ac depletion..... | 107 |
| Figure 5.7 Changes in gene expression and differentiation after IAA-induced KAT7 depletion in AID MOLM-13 and MV4-11 cells..... | 109 |
| Figure 5.8 MLL, AF9 and AF4 occupancy at KAT7-bound loci following KAT7 loss | 113 |

| | |
|---|-----|
| Figure 5.9 Co-localization of MLL-AF9, KAT7 and BRD4 at the promoter | 115 |
| Figure 5.10 Occupancy of BRD4 and RNA Polymerase II Serine 5 phosphorylation at KAT7-bound loci following IAA treatment | 116 |
| Figure 5.11 Changes in H3K14ac, H4K5ac and H4K8ac levels at KAT7-bound loci following depletion of KAT7 protein | 118 |
| Figure 5.12 Working model for KAT7 mechanism of action in regulating the expression of MLL-AF9 target genes..... | 119 |
| Figure 5.13 E508Q KAT7 mutant sensitizes MOLM-13 to IBET-151 treatment. | 120 |

List of Tables

| | |
|--|-----|
| Table 2.1 Cell lines used in this study and the oncogenic..... | 25 |
| Table 2.2 List of gRNAs used in this study..... | 42 |
| Table 2.3 Sequences of primers used for RT-qPCR..... | 43 |
| Table 2.4 Sequences of primers used for ChIP-qPCR..... | 43 |
| Table 2.5 List of antibodies used for ChIP and Western Blot | 44 |
| Table 5.1 Multiprotein complexes associated with MLL-fusion proteins | 98 |
| Table 5.2 Selective established MLL-fusion targets and the binding profiles of KAT7, MLL-AF9, MLL-AF4 and the earliest time point of gene down-regulation following KAT7 knock-out. | 110 |

List of Publication

Au, Y.Z., Gu, M, De Braekerleer, E., Yu, J.S., Ong, S.H., Gozdecka, M., *Chen, M., Tzelepis, K.*, Huntly, B.J.P, Yusa, K. & Vassiliou, G.S. KAT7 is a therapeutic vulnerability of MLL-rearranged acute myeloid leukemia [Status: submitted]

List of Abbreviations

| | |
|-----------------|--|
| Ab | antibody |
| AEP | AF4 family/ ENL family/ p-TEFb |
| AF10 | ALL1-Fused Gene From Chromosome 10 Protein |
| AF4 | ALL1-Fused Gene From Chromosome 4 Protein |
| AF6 | ALL1-Fused Gene From Chromosome 6 Protein |
| AF9 | MLLT3 Super Elongation Complex Subunit |
| AID | auxin inducible degron |
| AML | acute myeloid leukemia |
| B-ALL | B-cell acute lymphoblastic leukemia |
| BFP | blue fluorescent protein |
| bp | base pair |
| BRD4 | bromodomain containing protein 4. |
| BRPF | bromodomain and PHD finger-containing proteins |
| BSA | bovine serum albumin |
| CBP | CREB-binding protein |
| ChIP | chromatin immunoprecipitation |
| ChIP-seq | chromatin immunoprecipitation followed by DNA sequencing |
| CML | chronic myeloid leukemia |
| CpG | cytosine-phosphate-guanine |
| DGIdb | drug gene interaction database |
| DOT1L | disruptor of telomeric silencing 1- like histone lysine methyltransferase |
| DotCom | Dot1L complex |
| DSG | disuccinimidyl glutarate |
| E508Q | glutamic acid (E) to glutamine (Q) substitution at amino acid position 508 |
| EAP | elongation assisting protein/ENL associated proteins |
| EAP | ENL-associated proteins |
| EC | endothelial cell |
| EGS | ethylene glycol bis(succinic acid) |
| ELL | elongation factor for RNA polymerase II |
| ENL | MLLT1 Super Elongation Complex Subunit |
| FA | formaldehyde |

| | |
|----------------|--|
| FPKM | fragments per kilobase of transcript per million mapped reads |
| G485A | glycine (G) to alanine (A) substitution at amino acid position 485 |
| GFP | green fluorescent protein |
| gKAT7 | gRNA targeting KAT7 |
| gRNA | guide-RNA |
| gSIK3 | gRNA targeting SIK3 |
| h | hour |
| H3 | Histone 3 |
| H3K14ac | histone 3 lysine 14 acetylation |
| H3K4me3 | histone 3 lysine 4 trimethylation |
| H4 | Histone 4 |
| H4K12ac | histone 4 lysine 12 acetylation |
| HAT | histone acetyltransferase |
| HBO1 | histone acetyltransferase binding to ORC-1 |
| HDAC | histone deacetylase |
| HIPK1 | homeodomain interacting protein kinase 1 |
| HRP | horseradish peroxidase |
| IAA | indole-3-acetic acid |
| IP | immunoprecipitation |
| JADE | Jade Family PHD finger |
| KAT6A | lysine acetyltransferase 6A |
| KAT7 | lysine acetyltransferase 7 |
| KMD2A | lysine demethylase protein 2A |
| KMD5B | lysine demethylase protein 5B |
| LEDGF | lens epithelium-derived growth factor |
| MCM | minichromosome maintenance protein |
| MENIN | multiple endocrine neoplasia type 1 |
| Min | minutes |
| MLL | mixed lineage leukemia |
| MYST | Moz, Ybf2, Sas2p, Tip60 |
| p-TEFb | positive transcription elongation factor |
| PAF | polymerase associated factor |
| PAFc | polymerase-associated factor complex |
| PBS | phosphate-buffer saline |
| PFA | paraformaldehyde |

| | |
|------------------|---|
| PHD | plant homeodomain |
| PI | propidium iodide |
| Pol II S5 | RNA polymerase 2 serine 5 phosphorylation |
| RPKM | reads per kilobase of transcript per million mapped reads |
| RSK1 | ribosomal protein S6 kinase 1 |
| s | seconds |
| SEC | super elongation complex |
| Ser | Serine- rich |
| shRNA | short hairpin RNA |
| SIK3 | salt inducible kinase 3 |
| STK3 | serine/threonine kinase 3 |
| ZnF | zinc finger |

Chapter 1: Introduction

1.1 Acute myeloid leukaemia and disease pathogenesis

Acute myeloid leukaemia (AML) is an aggressive hematopoietic malignancy, characterised by the differentiation arrest and uncontrolled proliferation of myeloid progenitor cells (Döhner et al., 2015). Leukaemic blast cells infiltrate the peripheral blood, bone marrow and accumulate in other organs in the body. AML is associated with a plethora of clinical symptoms, including leukocytosis, anaemia, thrombocytopenia, bleeding and increased risk of infections (De Kouchkovsky and Abdul-Hay, 2016). The etiology of AML is not fully established, but it has been known that incidents increase with age (Bhayat et al., 2009; Dores et al., 2012). Other associated risk factors include genetic disorders and chemical or radiation exposure (Deschler and Lübbert, 2006).

The molecular pathogenesis have traditionally been studied using cytogenetic analysis, which elucidated many recurrent chromosomal structural abnormalities (Grimwade et al., 2015; Mrózek et al., 2004; Rowley, 2008). With next-generation sequencing technologies, recurrent gene mutations are increasingly apparent in AML and gene-gene interactions have been reported (Papaemmanuil et al., 2016; The Cancer Genome Atlas Research Network, 2013). Notably, the genetic aberrations found in AML can be classified into different functional categories, including transcription factor fusions (e.g. *PML-RARA*, *RUNX1-RUNX1T1*), tumour suppressors (e.g. *P53*, *WT1*), DNA methylation (e.g. *DNMT3A/B*, *TET1/2*, *IDH1/2*), activated signalling (e.g. *FLT3-ITD*, *KIT*, *KRAS/NRAS*), chromatin modifiers (e.g. *MLL*-fusions, *NUP98-NSD1*), myeloid transcription factors (e.g. *RUNX1*, *CEBPA*), cohesin, spliceosome and *NPM1* (The Cancer Genome Atlas Research Network, 2013). Beyond somatic mutations, dysregulated epigenetic mechanisms are thought to play a role in AML pathogenesis, as seen by the recurrent somatic mutations in genes encoding many epigenetic modifiers (Figueroa et al., 2010a; Rampal et al., 2014; Shih et al., 2015; The Cancer Genome Atlas Research Network, 2013). In addition, studies have reported alterations in genome-wide DNA methylation patterns in AML patients (Figueroa et al., 2010b; Li et al., 2016a; The Cancer Genome Atlas Research Network, 2013).

Despite the relatively low number of mutations per patient compared to other cancer types, genetic and epigenetic heterogeneity has been identified in AML (Lawrence et al., 2013, 2014; Li et al., 2016b). This heterogeneity can be found between different patients as well as within a patient. A founding leukaemia clone, which is defined by a clone with the highest variant allele frequency, and at least one

sub-clone are commonly detected in individual patients (The Cancer Genome Atlas Research Network, 2013; Welch et al., 2012). Understanding the underlying mutational profiles of AML can have implications on the therapeutic choice for the same patients at different stages of the disease (Grimwade et al., 2015).

1.2 Current and emerging treatments for AML

The conventional care for AML has not advanced for more than three decades, whereby induction chemotherapy serves as the front-line treatment for many patients (Döhner et al., 2016).

Anthracycline in combination with cytarabine are given to patients suitable for intense chemotherapy, which include those younger than 60 years of age or older but showing favourable- or intermediate-risk cytogenetics (Döhner et al., 2015, 2016). Anthracycline is an antibiotic produced by *Streptomyces peucetius varcaesitue* and thought to interact with nuclear components such as DNA and topoisomerase, causing DNA damages (Minotti, 2004; Rabbani et al., 2005). Cytarabine is a pyrimidine analogue and the mechanism of cytotoxicity is presumably due to the cumulative effect of DNA polymerase inhibition and competitive incorporation of the nucleoside analogue into DNA; the latter results in chain termination and subsequent block of DNA synthesis (Galmarini et al., 2001; Kufe et al., 1980; Major et al., 1981, 1982). Post-remission treatments, also known as consolidation therapy, include further chemotherapy and hematopoietic-cell transplantation. These treatments aim to eliminate any residual diseased cells and prevent relapse (De Kouchkovsky and Abdul-Hay, 2016).

Although the majority of the patients younger than the age of 60 achieve complete remission after induction therapy, the response rates are less optimistic in older AML patients (Döhner et al., 2015; Lowenberg et al., 1999). Relapse and subsequent resistance to chemotherapy are major barriers to the long-term survival of AML patients. It is estimated that only about 30% of the patients are cured of the disease, whilst the majority of the patients experience relapse (Döhner et al., 2015, 2016; Liesveld, 2012). This relapse of the disease is associated with the clonal evolution of leukaemic blasts within individual patients (Garson et al., 1989; Testa et al., 1979). A recent study using next generation sequencing has revealed the pattern of clonal evolution of AML cells that led to disease relapse (Ding et al., 2012). The authors found that during or after chemotherapy, primary AML clones acquired additional mutations and gave rise to relapse clones. They suggested that although essential for initial remission, chemotherapy could introduce relapse-specific mutations into founding AML clones or surviving sub-clones, which can undergo selection and clonal expansion.

The effort to increase treatment options for AML patients can be reflected by the numerous novel agents that are currently in different phases of development. New formulation of chemotherapies includes topoisomerase inhibitors (e.g. Vosaroxin), hypomethylating agents (e.g. Guadecitabine) and CPX-351. CPX-351 combines cytarabine and anthracyclin daunorubicin in 5:1 molar ratio and is encapsulated in liposome (Döhner et al., 2015; Stein and Tallman, 2015). In addition, an antibody-drug conjugate (ADC) was developed to deliver the conjugated drug specifically to CD33⁺ leukaemic cells and have shown myelosuppressive effects (Döhner et al., 2015; Stein and Tallman, 2015). Many molecular targeted inhibitors are also under evaluation in clinical trials. Agents that target recurrent mutations include inhibitors against FLT3 (e.g. Quizartinib, Crenolanib, Gilteritinib) (Döhner et al., 2015; De Kouchkovsky and Abdul-Hay, 2016; Stein and Tallman, 2015), IDH1/2 (e.g. AG-221, AG-120) (Döhner et al., 2015; Pollyea et al., 2014; Stein et al., 2014) and BCL2 (DiNardo et al., 2019; Kontro et al., 2015; Pollyea et al., 2018; Pullarkat and Newman, 2016). DNA methyltransferase (DNMT) inhibitors Azacitidine and Decitabine have been approved for clinical use (Wouters and Delwel, 2016). Other epigenetic modifier inhibitors in clinical trial include a histone lysine methyltransferase DOT1L inhibitor EPZ-5676, a histone lysine demethylase LSD1 inhibitor GSK2879552, histone deacetylase (HDAC) inhibitors Pracinostat, Vorinostat and Panobinostat, and Bromodomain (BET) inhibitor OTX-015 (Daigle et al., 2011; Döhner et al., 2015; Grimwade et al., 2015; Stein and Tallman, 2015; Wouters and Delwel, 2016). Potential targets in preclinical development include histone lysine methyltransferase inhibitors against EZH2 (Girard et al., 2014; Knutson et al., 2012), MLL-Menin (Borkin et al., 2015; Grembecka et al., 2012) or MLL-LEDGF (Čermaková et al., 2014) interface, histone lysine acetyltransferase inhibitors against CREBBP (Gang et al., 2014) and EP300 (Gao et al., 2013b), as well as histone arginine methyltransferase PRMT inhibitors (Alinari et al., 2015; Chan-Penebre et al., 2015). Notably, the emergence of many epigenetic targeted therapies in AML reflects the involvement of epigenetics in the pathogenesis of AML.

1.3 CRISPR-Cas9 as a powerful tool to identify vulnerabilities in AML

Understanding of the genetic contribution to diseases has been enhanced by advances in targeted gene-editing technologies, such as the clustered regularly interspaced short palindromic (CRISPR)-Cas (CRISPR-associated) system. Originally a defence mechanism found in bacteria and archaea (Horvath and Barrangou, 2010), the CRISPR-Cas system has been exploited in a wide range of organisms, including mammalian cells (Bassett et al., 2015; Chen et al., 2015; Hart et al., 2015; Matano et al., 2015; Sánchez-Rivera and Jacks, 2015; Shi et al., 2015). This system composes of Cas protein and CRISPR RNA (crRNA). The latter possesses nucleotide sequence complementary to target region(s) of the genomic DNA. Depending on the types of CRISPR system (I-III), the manner in which precursor

CRISPR RNA (pre-crRNA) is processed into mature crRNA and the associated effector Cas protein differ.

Type II CRISPR-Cas9 system is one of the best characterised versions and most commonly used for genome engineering due to its simplicity in terms of the number of genes involved (Makarova et al., 2015). While type I and III CRISPR-Cas systems utilise a large multimeric crRNA-Cas complex to achieve nuclease activity, the type II system requires a single Cas9 endonuclease with crRNA and trans-activating crRNA (tracrRNA) and cleave target DNA sequences (Makarova et al., 2011). In the engineered CRISPR-Cas9 system, crRNA and tracrRNA are fused together and the system requires single guide-RNA (gRNA) (Ran et al., 2013). The simplicity has accelerated its development as a molecular tool (Rath et al., 2015). Specificity towards a particular region of the genome relies on a 20-nucleotide variable region at the 5' end of the gRNAs (Koike-Yusa et al., 2014). Double stranded break (DSB) triggers DNA repair pathways such as non-homologous end joining (NHEJ) or homology-directed repair (HDR). Exploiting the inaccurate repair of DNA by NHEJ, nucleotide insertion or deletion (indels) can lead to frameshift mutations that derives non-functional protein (Ran et al., 2013; Rath et al., 2015; Sander and Joung, 2014). Compared to RNA interference (RNAi), which utilises small interfering RNA (siRNA) or short hairpin RNA (shRNA) to generate knock-down of target genes, CRISPR-Cas9 creates targeted gene knock-out using gRNA. Both technologies suffer disadvantages. RNAi often results in incomplete suppression of gene expression, whereas the use of CRISPR-Cas9 is restricted to target regions in the genome with protospacer adjacent motifs (PAM). The possibility of off-targets is common to both technologies (Koike-Yusa et al., 2014; Ran et al., 2013; Sander and Joung, 2014; Shalem et al., 2015). Nevertheless, CRISPR-Cas9 serves as an invaluable tool to address many biological questions.

The ability to generate bi-allelic mutations at a targeted locus has sparked many questions in the field of cancer research (Chen et al., 2015; Hart et al., 2015; Sánchez-Rivera and Jacks, 2015; Shi et al., 2015). In particular, genome-wide loss-of-function genetic screening allows the identification of recessive genes important in cancer. CRISPR-Cas9 screens have been performed in human AML cell lines (Erb et al., 2017; Tzelepis et al., 2016; Wang et al., 2017). This method elucidated essential genes that drive leukaemic cells. Utilising AML cell lines with different genotypes informed genotype-specific essentiality. In the CRISPR screen performed by Tzelepis *et al.*, they identified 492 AML-specific essential genes across 5 AML cell lines (MOLM-13, MV4-11, HL-60, OCI-AML2 and OCI-AML3), 227 of which are druggable, defined as proteins with enzymatic activity which can potentially be targeted (Tzelepis et al., 2016; Wagner et al., 2016). Potential novel therapeutic targets within these large datasets may be teased out to ameliorate the current treatment landscape of AML.

It has been proposed that targeting leukaemia-specific mutations or dependencies is one of the successful examples to achieve higher cure rates in patients (Liesveld, 2012). Given that cure remains elusive for the majority of the AML population and the conventional treatment regimens for AML is non-specific, toxic and unsuitable for high-risk patients, this thesis aims to identify novel and promising druggable candidates by understanding the mechanism of their essentiality. Genome-wide CRISPR-Cas9 dropout screens revealed KAT7 as a novel AML-specific vulnerability that could potentially be exploited therapeutically (Tzelepis et al., 2016).

1.4 KAT7 protein complexes and histone tail specificity

KAT7, also known as MYST2 or HBO1, is a histone lysine acetyltransferase (HATs/KATs) belonging in the MYST family proteins (Avvakumov and Côté, 2007; Sheikh and Akhtar, 2018; Voss and Thomas, 2009). The defining and common feature of MYST family protein is the highly conserved MYST catalytic domain, containing an acetyl-CoA binding motif and a C2HC-zinc finger motif (Avvakumov and Côté, 2007; Voss and Thomas, 2009). Other members of the MYST family include KAT5 (TIP60), KAT6A (also called MOZ or MYST3) and KAT6B (MORF or MYST4). Possessing only the zinc finger, serine-rich and MYST domains, KAT7 relies on forming multi-subunit complexes with scaffold protein JADE or BRPF, along with Esa1 associated factor 6 (EAF6) and inhibitor of growth protein 4/5 (ING4/5) for chromatin binding (Figure 1). ING4/5, JADE and BRPF subunits all possess the plant hormone domain (PHD), which is a recognition motif that binds to histone modifications on the N-terminal tail of histone H3 (Avvakumov et al., 2012; Lalonde et al., 2013; Musselman et al., 2012; Saksouk et al., 2009).

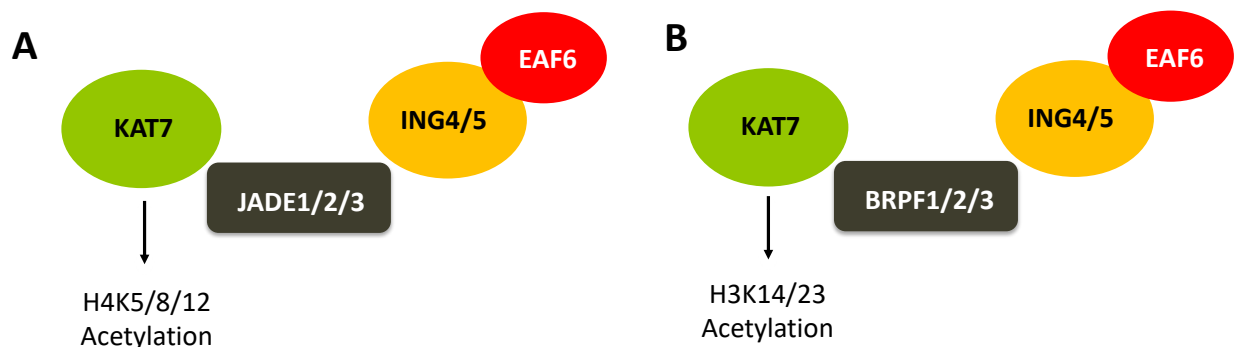


Figure 1 Protein complexes of KAT7

KAT7 is found in complexes containing ING4/5, EAF6 and JADE1/2/3 or BRPF1/2/3. A) KAT7-JADE complex facilitates histone H4 acetylation on lysine residues 5, 8 and 12. B) KAT7-BRPF complex has affinity for histone H3 tails, acetylating at lysine residues 14 and 23.

ING4/5 consist of one PHD domain while both JADE and BRPF each possess two PHD domains. KAT7 complex is localized to H3K4me3 enriched region of the chromatin through the PHD domain of ING4/5 (Lalonde et al., 2013; Saksouk et al., 2009). However, the histone tail specificity of KAT7 is determined by the scaffold protein in the complex— KAT7-BRPF and KAT7-JADE complex has an affinity to histone 3 tails (H3K14ac, H3K23ac) (Feng et al., 2016; Kueh et al., 2011; Lalonde et al., 2013; Mishima et al., 2011) and histone 4 tails (H4K5ac, H4K8ac, H4K12ac) (Doyon et al., 2006; Iizuka et al., 2009) respectively.

All paralogues of JADE and BRPF proteins have multiple shared properties. Firstly, the two PHD domains are linked by zinc knuckle and are known as the PZP domain that is essential for chromatin binding (Avvakumov et al., 2012; Lalonde et al., 2013). Secondly, flanking the PZP domains are domain I and domain II where the former interacts with KAT7 and the latter binds ING4/5 and EAF6. It is hypothesized that the PZP domain of the scaffold protein drives the acetylation of lysine residues on H3 tail, however, the N-terminal region (EPcA-related N-terminal domain) just before the domain I in JADE protein is crucial for the histone H4 tail specificity and lysine acetylation on H4 tails (Lalonde et al., 2013).

1.5 Molecular, cellular and biological functions of KAT7

1.5.1 DNA-replication

KAT7 is involved in the initiation of DNA replication by regulating the assembly/formation of the pre-replication complex (pre-RC) (Doyon et al., 2006; Iizuka et al., 2006; Miotto and Struhl, 2008, 2010, 2011). In G1 phase of the cell cycle, KAT7 is recruited to the replication origins by direct binding with Cdt1 and serving as a co-activator of Cdt1, consequently affect the licensing activity of Cdt1 (Miotto and Struhl, 2008). The Cdt1-dependent recruitment of KAT7 leads to global histone 4 acetylation at the origin, which is positively correlated with the loading of minichromosome maintenance (MCM) protein complex required for the firing of replication origins (Miotto and Struhl, 2010). Potential involvement of non-histone acetylation by KAT7 in controlling this process has been speculated due to in vitro evidence of KAT7 acetylating subunits of the pre-RC such as Orc2, Geminin, Mcm2 and Cdc6 (Iizuka et al., 2006). Furthermore, KAT7 has also been reported to play a role in the S phase of the cell cycle for DNA synthesis (Doyon et al., 2006). Together, these evidences illustrate the catalytic-dependent activity of KAT7 in regulating the cell cycle.

1.5.2 Transcription and regulation of gene expression

KAT7 has been implicated in transcription control in various cellular contexts, in part, via its histone acetylation function. Occupancy of KAT7 at the 5' region of genes positively correlated with gene expression in RKO colon carcinoma cell lines, with enrichment of KAT7 found both upstream and downstream of TSS (Avvakumov et al., 2012). In addition to promoter binding, KAT7 is localized at/to gene bodies. Tetrameric KAT7 complex (KAT7-JADE-ING4/5-EAF6) localizes to H3K4me3-enriched chromatin region to acetylate lysine residues on histone tails, whilst in dimeric form, KAT7-JADE has affinity to H3K36me3, a modification that marks active coding regions of the gene (Barth and Imhof, 2010; Wu and Snyder, 2008), in order to deposit H4 acetylations across the gene body (Saksouk et al., 2009). This suggests that KAT7 plays a role not only in the initiation stage but also in the elongation of transcription. Furthermore, intragenic region (also known as intronic region) H3K14ac and pan-H4 acetylations by KAT7 is implicated in regulating the transcription of endothelial genes such as VEGFR-2 transcription in human and zebrafish endothelial cells (Yan et al., 2018).

Several studies suggested that KAT7 could modulate the transcriptional programme through functions other than histone acetylations. In particular, multiple studies have associated KAT7 as co-factors of hormones receptors (Georgiakaki et al., 2006; Iizuka et al., 2013; IIZUKA et al., 2017; Sharma et al., 2000). Intrinsic E3 ligase activity of the MYST domain of KAT7 has been linked to the ubiquitination of estrogen receptor (ER) which destabilizes the ER for proteasomal degradation (Iizuka et al., 2013; IIZUKA et al., 2017). In another study, Sharma et al., identified interactions between KAT7 and androgen receptor (AR) both in vitro and in vivo (CV-1 monkey kidney cell line). Binding of KAT7 to AR represses AR-mediated transcription and this is thought to be mediated via the N-terminal region of KAT7 (Sharma et al., 2000). In contrast to the findings of Iizuka et al., Sharma et al, did not detect inhibitory effect of KAT7 on ER-dependent transcription. KAT7 is a co-activator of various other nuclear hormone receptor, including progesterone receptor to modulate its transcriptional activity (Georgiakaki et al., 2006). Importantly, KAT7 may regulate transcription by physical interaction with nuclear factors and/or via its non-histone acetylation function. Depending on the context, KAT7 may promote or suppress gene expression.

1.5.3 Physiological roles of KAT7

Loss of KAT7 has severe consequences in development— mice embryos lacking KAT7 do not develop beyond mouse embryonic day 10.5 (E10.5) and is arrested at the 10-somite stage due to growth

retardation and down regulation of genes essential for embryonic patterning (Kueh et al., 2011). Importantly, KAT7 is responsible for global H3K14ac in mouse embryos and primary embryonic fibroblasts and the KAT7 mutant of these cell types were able to proliferate normally and showed no defects on DNA replication in vitro, suggesting KAT7 is dispensable for general DNA replication, cell cycle and proliferation (Kueh et al., 2011). KAT7 is also required to maintain global H3K14ac in fetal livers and for the expression erythroid regulator genes in erythropoiesis (Mishima et al., 2011). Together, these studies imply the critical role of KAT7 as a transcription regulator for important developmental genes.

1.6 KAT7 in cancer and diseases

The role of KAT7 in cancer is not well explored. A study has found the concentration of KAT7 protein in cancer cell lines are more abundant compared to normal cells (Iizuka et al., 2009). Among primary cancers, tissues from testis, breast, ovary and bladder tumours have high KAT7 protein levels (Chen et al., 2017; Iizuka et al., 2009). Knock-down of KAT7 has anti-tumour effects in bladder cancer cells, via the Wnt/B-catenin signalling pathway, and has been proposed as a potential therapeutic target in the treatment of this carcinogenesis (Chen et al., 2017). Although KAT7 is also normally highly expressed in testis (Iizuka et al., 2009; Sharma et al., 2000), it possesses putative tumour suppressive roles in hormone receptor driven tumours such as prostate and breast cancers (Iizuka et al., 2013; IIZUKA et al., 2017; Sharma et al., 2000). Depending on the context, HATs can serve as an oncogene or tumour suppressors. In AML, HATs of the MYST family are frequent targets of chromosomal translocations (Sheikh and Akhtar, 2018). While KAT6A and KAT6B are known targets of chromosomal translocations that drive AML (Borrow et al., 1996; Carapeti et al., 1998; Esteyries et al., 2008; Kitabayashi et al., 2001), the role of KAT7 in AML, if any, remains poorly understood.

1.7 Aims

This thesis aims to validate dropout hits from the genome-wide AML CRISPR-Cas9 screens and further elucidate the molecular mechanism of essentiality of the most promising candidate. In particular, to decipher the potential role of KAT7 in the pathogenesis of AML, by addressing the following broad objectives:

- Investigate the phenotypic consequences of KAT7 *knock-out* in various AML cell lines.
- Characterize the molecular functions of KAT7 in cellular models of AML.
- Understand the mechanism of action of KAT7 in vulnerable AML subtypes.

Chapter 2: Materials and methods

2.1 Cell culture

All Cas9-expressing cell lines, with the exception of Nomo-1, were kindly provided by the Vassiliou group. Cas9-expressing MOLM-13, MV4-11, OCI-AML2 and OCI-AML3 were generated in the study performed by Tzelepis et al. Cas9-expressing THP-1 and K562 cell lines were generated in a separate study in Vassiliou's group. Nomo-1 cell line was derived from Mathew Garnett. Generation of Cas9 expressing Nomo-1 is described below.

MOLM-13, MV4-11, THP-1, HL-60 and K562 were cultured in RPMI 1640 medium (52400-25, Life Technologies) supplemented with 10% fetal bovine serum (10270-106, Life Technologies). OCI-AML2 and OCI-AML3 were cultured in MEM- α (BE12-169-F, Lonza) with 20% fetal bovine serum (10270-106, Life Technologies) and 1% GlutaMax (35050-038, Life Technologies). Nomo-1 was cultured in RPMI 1640 supplemented with 10% FBS, 1% penicillin/streptomycin (15070063, Life Technologies), 1% sodium pyruvate (11360039, Life Technologies), 1% glucose (Sigma). To maintain Cas9 expression, blasticidin (R210-01, Invitrogen) was added to MOLM-13, MV4-11, THP-1, Nomo-1, HL-60, OCI-AML2 and OCI-AML3 at 10 $\mu\text{g/ml}$, and K562 at 15 $\mu\text{g/ml}$. 293 FT cells were cultured in DMEM (21969-035, Life Technologies) supplemented with 10% fetal bovine serum (10270-106, Life Technologies) and 1% GlutaMax (35050-038, Life Technologies). All cell lines were incubated at 37 °C with 5% CO₂. The oncogenic mutations and patient profile of cell lines used in this study are indicated in Table 2.1.

| | Cell line | Oncogenic mutations |
|----------------|-----------|--------------------------------------|
| MLL-fusion | MOLM-13 | <i>MLL-AF9, FLT3-ITD</i> |
| | MV4-11 | <i>MLL-AF4, FLT3-ITD</i> |
| | OCI-AML2 | <i>MLL-AF6, DNMT3A</i> |
| | THP-1 | <i>MLL-AF9, TP53</i> |
| | Nomo-1 | <i>MLL-AF9, KRAS</i> |
| Non MLL-fusion | OCI-AML3 | <i>NRAS, NPM1, BAX, DNMT3A</i> |
| | HL-60 | <i>NRAS, TP53, CDKN2A</i> |
| | K562 | <i>BCR-ABL, TP53, CDKN2A, PDGFRA</i> |

Table 2.1 Cell lines used in this study and the oncogenic

2.2 gRNA design and cloning of gRNA into expression vector

The two best-performing gRNAs (unpublished data) from the genome-wide CRISPR-library used in the initial screens (Tzelepis et al., 2016) were selected for the initial validation of candidate genes. For *KAT7* and *SIK3*, additional 4-5 gRNAs were designed using WTSI Genome Editing (WGE) (<http://www.sanger.ac.uk/htgt/wge/>). Only gRNAs that target the exon of all putative transcripts and have no less than 3 nucleotide off-target mismatches in the sequence were selected. All gRNAs were cloned into the BbsI site of pKLV2-U6gRNA5(insert)-PGKpuro2ABFP-W (Figure 2.1D) (Koike-Yusa et al., 2014).

2.3 Virus production and transduction

Lentivirus was produced by transfecting 293 FT cells in one well of 6-well plates. 0.9 µg lentivirus vector, 0.9 µg psPax2 packaging vector (12260, Addgene), 0.2 µg pMG2.G (12259, Addgene) and 2 µl of PLUS reagent (15338100, Invitrogen) were added to 500 µl Opti-MEM (51985026, Invitrogen) and mixed by vortex. The mixture was incubated for 5 min at room temperature. 6 µl of Lipofectamine LTX (15338100, Invitrogen) were then added to the mixture, vortexed and further incubated for 30 min at room temperature. The transfection mixture was added to each well and incubated for 6-8 h at 37 °C with 5% CO₂ and then replaced with DMEM (21969, Invitrogen) supplemented with 10% fetal bovine serum (10270-106, Life Technologies) and 1% GlutaMax (35050-038, Life Technologies). Viral supernatant was collected on day 2, filtered with 0.45µm surfactant-free cellulose acetate (SFCA) syringe filter (190-2545, Nalgene) and stored at -80 °C.

AML and CML cell lines were transduced by adding lentivirus and 8 µg/ml polybrene (Millipore) to 3.0 x 10⁴ cells per well of 96-well plate (for proliferation assay) or 1 x 10⁶ cells per well of 6-well plate and incubated for 22 h at 37 °C. Viral supernatant was then replaced with culture media and cells were passaged. BFP-positive cells were sorted on day 3 post-transduction and used for subsequent functional assay (see below).

2.4 Generating Cas9 expressing Nomo-1 cell line and blasticidin selection

Cas9 construct is integrated into the genome by lentivirus transduction. The Cas9 expression vector is illustrated in Figure 2.1A and previously reported by Tzelepis et al. The concentration required for blasticidin selection was determined by a “kill curve” assay, which is the minimum concentration of

blasticidin to kill all wild-type cells. Concentration tested: 1, 5, 10, 15, 30 $\mu\text{g/mL}$. Blasticidin concentration $>10 \mu\text{g/mL}$ killed all Nomo-1 cells.

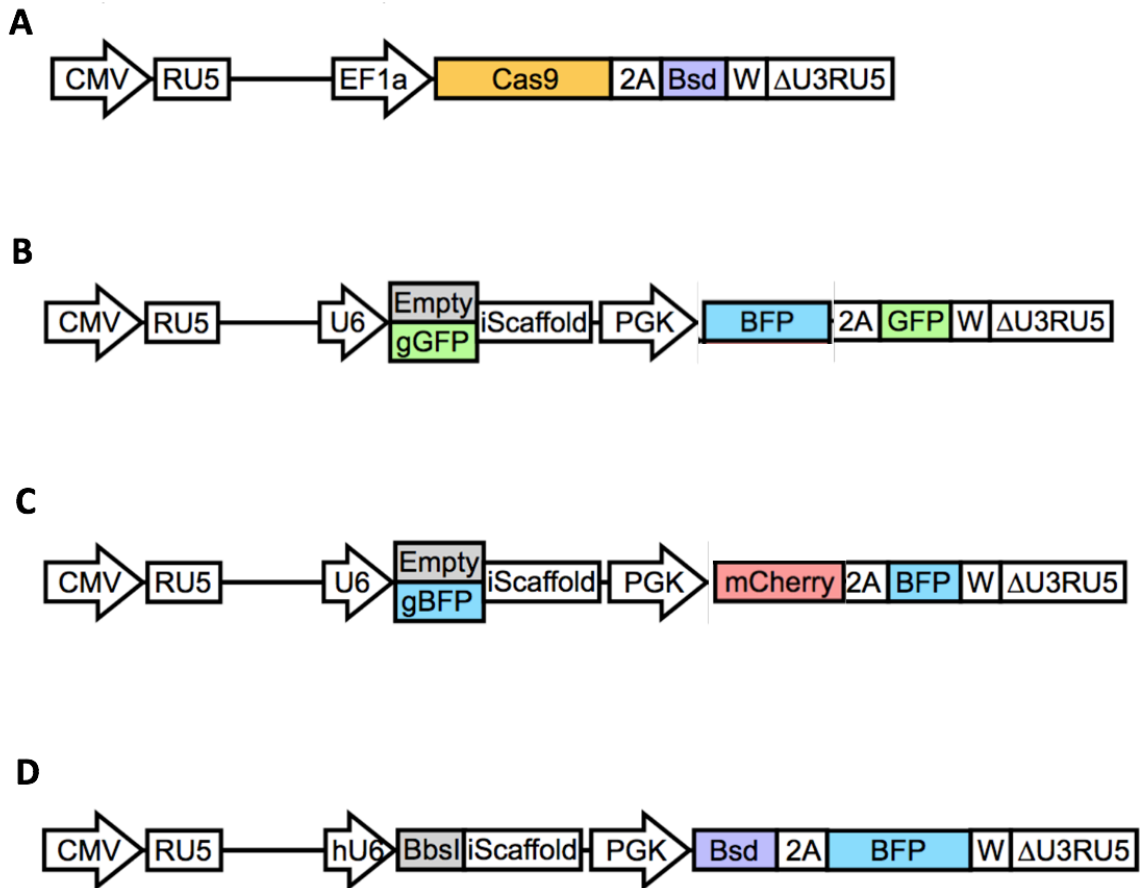


Figure 2.1 Cas9, Cas9 reporter and gRNA constructs

A) The Cas9 construct used to generate the Cas9 expressing cell lines used in this study. The Cas9 gene and blasticidin resistant (Bsd) genes are expressed under the EF1-alpha promoter. B-C) Cas9 reporter construct to assess the Cas9 activity of cells. Gene encoding fluorescence proteins are expressed with the PGK promoter and gRNA are expressed under the U6 promoter. B) BFP and GFP expression with gRNA targeting GFP (gGFP). C) mCherry and BFP expression with gRNA targeting BFP (gBFP). D) gRNA scaffold construct with gRNA insert at the BbsI restriction site, expressed under the U6 promoter. Selectable markers Bsd and BFP are expressed under the PGK promoter.

2.5 Cell sorting and FACS analyses

Cell sorting was performed using Mo-Flo XDP or BD INFLUX under containment-level 2 (CL2) condition. FACS analyses for functional assays were performed using BD LSRFortessa. All raw files were analysed using FlowJo.

2.6 Sub-cloning and characterisation of Cas9-expressing AML cells

Single cells of MOLM-13, MV4-11, THP-1, HL-60, OCI-AML2, OCI-AML3 and MOLM-13^{luciferase} were sorted into individual wells of 96-well plate and cultured for 2-3 weeks until colonies were observed. Growing colonies were transferred to 24-well then 12-well plates, further expanded. Cas9 activity was examined using the Cas9 reporter construct pKLV2-U6gRNA (target GFP)-PGK-BFP-T2A-GFP (Figure 2.1B) (Tzelepis et al., 2016). As the presence of luciferase protein in MOLM-13^{luciferase} give rise to autofluorescence that interferes with the GFP channel, we opted to use a variant reporter whereby GFP is replaced by BFP and BFP is replaced with mCherry: pKLV2-U6gRNA (target BFP)-PGK-mCherry-BFP-mCherry (Figure 2.1C) (Tzelepis et al., 2016) to assess the Cas9 activity of this cell line.

2.7 Functional assays

2.7.1 Proliferation

The validation of each candidate gene was performed in triplicates in 96-well plates. The relative percentage of BFP-positive and BFP-negative cells was determined for each well every 2 days between day 4 and 12 post-transduction by FACS. Culture media were refreshed every two days. Cells were fixed in 4% Paraformaldehyde (PFA) in phosphate-buffered saline (PBS) for 10 min and resuspended in 1% bovine serum albumin (BSA) in PBS before analysis.

2.7.2 Differentiation

2.3×10^5 cells were harvested to stain for CD11b (17-0118, eBioscience) or CD13 (17-0138, eBioscience). APC-conjugated mouse IgG1 κ isotype control (17-4714, eBioscience) was used to detect non-specific binding. Cells were washed once in PBS and once in staining buffer (2% fetal bovine serum in PBS), followed by resuspension in anti-human CD11b or CD13 antibody solution

(1.15 µl antibody per 100 µl staining buffer per test) for 30 min on ice in the dark. After incubation, cells were washed twice with staining buffer and then analysed by FACS.

2.7.3 Apoptosis

2.3×10^5 cells were analysed using Annexin V Apoptosis Detection Kit APC (88-8007, eBioscience). Cells were washed once with PBS and once in 1X binding buffer. 10X binding buffer (88-8007, eBioscience) was diluted by adding 1 mL 10X binding buffer with 9 mL dH₂O. Cell pellets were resuspended in 100 µL annexin V binding buffer (1.15 µL Annexin V-APC per 100 µL 1X binding buffer) and incubated in dark for 15 min in room temperature. Next, cells were washed in 1X binding buffer and resuspended in 200 µL of the same buffer. 1.15 µL of PI solution (88-8007, eBioscience) was added immediately before flow cytometry analysis. Samples are kept on ice following staining.

2.7.4 Cell cycle

2.3×10^5 cells were washed twice with PBS and fixed in ice-cold 70% ethanol for at least 24 h at 4 °C. Fixed cells were washed once in PBS and resuspended in Propidium Iodide (PI) staining solution [Triton X-100 (0.1% v/v), RNase A (50 µg/ml) and PI (25 µg/ml)] overnight at 4 °C. DNA content was analysed by flow cytometry.

2.8 Cloning and generation of gRNA-resistant wild-type, G485A and E508Q mutant KAT7 cell lines

KAT7 cDNA (wild-type and G485A) that is resistant to gKAT7 (5) and gKAT7 (A10) targeting was cloned into a separate lentiviral plasmid backbone (EF1a-GFP-2A). Briefly, the plasmid was linearized by BsrGI (NEB) and NotI (NEB). KAT7 cDNA, in the form of gBlock, is subsequently cloned by Gibson assembly according to the manufacturer's guidelines. The resulting plasmid pKLV-EF1aGFP2AKAT7-W, expressing wild-type KAT7 was used to generate E508Q transgene. The KAT7-plasmid was double digested sequentially by BsaBI (NEB) followed by SrfI (NEB) to remove the specific segment in the MYST region of the wild-type KAT7 sequence. Sequential digestion was used as the optimal temperature for enzymatic activity differs for these two restriction enzymes. The volume of enzyme added is typically <10% of reaction volume where reaction volume is up to 50 µL.

The plasmid is first digested by BsaBI in following mixture and conditions:

| | | |
|--------------------|---------|---|
| 0.5 µg Plasmid DNA | 2.5 µL | <div style="display: inline-block; vertical-align: middle; text-align: center;"> <div style="border-left: 1px solid black; padding-left: 10px; margin: 0;"> 60 °C for 90 min 80 °C for 20 min Store at 4 °C </div> </div> |
| BsaBI | 2.5 µL | |
| CutSmart | 37.5 µL | |
| dH ₂ O | 47.5 µL | |
| | 47.5 µL | |

The BsaBI digested plasmid DNA from the above reaction was then digested by SrfI as follows:

| | | |
|--------------------|---------|--|
| BsaBI digested DNA | 47.5 µL | <div style="display: inline-block; vertical-align: middle; text-align: center;"> <div style="border-left: 1px solid black; padding-left: 10px; margin: 0;"> 37 °C for 2 h Store at 4 °C </div> </div> |
| SrfI | 2.5 µL | |
| | 50 µL | |

Linearized DNA was extracted by first running the digested DNA in 1% agarose gel (Figure 2.2) and using the QIAquick Gel Extraction kit (28704, Qiagen), following manufacturer's protocol. The image of the gel (Figure 2.2) shows the linearized DNA and the circular DNA where the latter appears to be smaller in length. This is because circular DNA migrates more rapidly than the linearized plasmid of the same mass. Furthermore, a faint band of 600 bp can be seen and this corresponds to the other digested product that is smaller in length. Both of these confirms that the plasmid DNA has been digested.

Next, we used the Gibson assembly to ligate the linearized vector with E508Q gBlock. Gibson assembly was performed following the manufacturer's protocol (E2611, NEB). The plasmid was amplified by transforming NEB 5-alpha Competent *E. coli* (C2987H/I, NEB) and plated on ampicillin agar plates to select for plasmids, conferring ampicillin resistant, that are successful ligated. The next day, colonies were picked and grown in 2xTY media + 50 µg/mL ampicillin. The plasmids were purified from *E. coli* using Miniprep Kit (27104, Qiagen). Cloning was confirmed by Sanger sequencing. The transgene was introduced to MOLM-13 cells via lentiviral transduction.

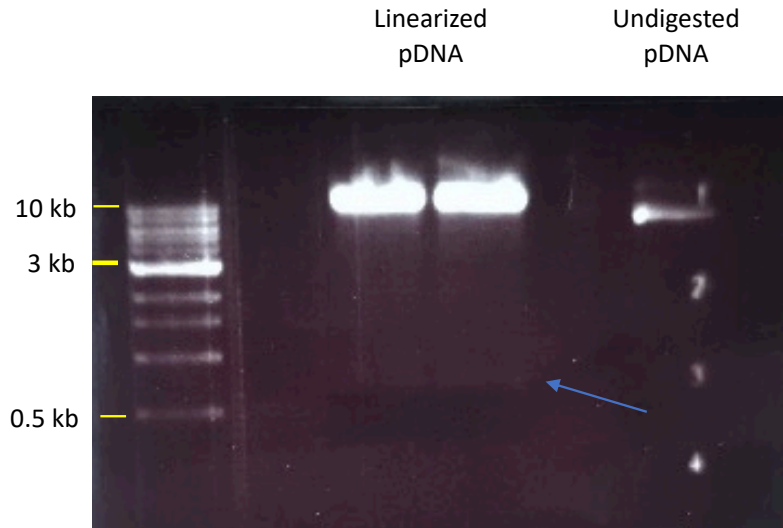


Figure 2.2 Double digestion of plasmid backbone for the integration of E508Q mutant sequence

The plasmid pKLV-EF1aGFP2AKAT7-W was digested with BsaBI and SrfI to linearize the plasmid backbone and for the cloning of the sequence corresponding to the E508Q mutation by Gibson assembly. The thick bands in lane 3 and 4 are the linearized plasmid DNA (pDNA). The blue arrow indicates a faint band at ~600bp, corresponding to the digested product. Note that the undigested pDNA in lane 6 migrate faster along the agarose gel, compared to the linearized pDNA.

2.9 Sanger Sequencing

LIGHTRun tube NightXpress service was used for capillary sequencing to identify successful cloning. Samples were prepared as follows: 5 μ L of purified plasmid DNA at a concentrate of 80-100 ng/ μ L mixed with 5 μ L of 5 μ M primer.

2.10 Western blot

The total number of cells was quantified by TC20 cell counter (Bio-Rad). Cells were harvested by centrifugation and washed twice in PBS. Cell pellets were then resuspended in NuPAGE LDS sample buffer (NP0007, ThermoFisher), NuPAGE sample reducing agent (NP0009, ThermoFisher) and water at a concentration of 1×10^6 cells/ 100 μ L. Samples were heated at 95 $^{\circ}$ C for 5 min and vortexed at room temperature for 10 min. An equal volume of samples was loaded into each well of Bis-Tris gels and electrophoresis were run at 150 V for 60 min in 1X MOPS SDS running buffer (NP0001, ThermoFisher). Polyvinylidene difluoride (PVDF) blotting membrane (10600023, Amersham) was

activated by suspending in 100% methanol for 2 min and washed in water before suspending in 1X transfer buffer (NP0006-1, ThermoFisher). Gel to membrane transfer was performed in ice-cold 1X transfer buffer at 90-110 V for 1.5 h or until good separation. Membrane was blocked with 5% milk (170-6406, Bio-Rad) or BSA (A9647, Sigma) in 0.1% TBS-T (1X TBS buffer (28358, ThermoScientific) with 0.1% *Tween* 20 (P9416, Sigma)) for at least 30 min. Primary antibodies were incubated at 4 °C over-night with rolling or shaking. Membranes were washed 3 times in 0.1% TBS-T for 10 min each the following day. The secondary antibody was blocked in 5% milk or BSA at room temperature for 1 h with rolling or shaking. Lastly, membranes were washed 3 times in 0.1% TBS-T for 10 min each. Enhanced chemiluminescent (ECL) (RPN3244, Amersham) or SuperSignal Pico PLUS (34577, ThermoScientific) was used as the substrate for the detection of horseradish peroxidase (HRP) on immunoblots.

All antibodies used for western blots are listed in Table 2.5.

2.11 In vivo mouse work

MOLM-13 luciferase-expressing cells were transduced with lentivirus for gRNA-mediated knock-out of KAT7 using gKAT7 (A10) or empty control. BFP-positive cells were sorted at day post-transduction. 5×10^5 cells were injected via the tail vein of immunocompromised NSGW41 male mice (derived by breeding NSG and KITW41 animal to homozygosity, i.e. *Kit*^{W41/W41}, *Prkdc*^{-/-}, and *Il2rg*^{-/-} or *Il2rg*^{+/-}) on day 3 post-transduction. 5 mice were injected in each treatment group.

Bioluminescence was done by injection of 2 mg/mL of D-luciferin (BioVision) per mouse. Mice quantification of bioluminescence was performed using IVIS Spectrum In Vivo Imaging System (PerkinElmer), with Living Image version 4.3.1 software (PerkinElmer) according to the manufacturer's instructions. Log-rank (Mantel-Cox) test was performed to measure statistical significance.

2.12 Reverse transcription-quantitative polymerase chain reaction (RT-qPCR)

For each sample, 500 ng RNA was made up to a volume of 11.875 μ L in nuclease-free water. Then, 0.5 μ L random primers (C1181, Promega) and 1 μ L dNTP (U1511, Promega) were added to each RNA sample. Next, the samples were incubated at 65 °C for 5 min then immediately snap cool on ice for 5 min. After RNA and primers were denatured, RNA was reverse transcribed to cDNA by adding 4 μ L 5x first-strand buffer (Invitrogen, 18064-071), 2 μ L 0.1M DTT (18064-071, Invitrogen), 0.5 μ L RNase

out (10777-019, Invitrogen) and 0.125 μ L Superscript II (Invitrogen, 18064-071) to each sample. Then, the mixture was placed in PCR machine with the following cycles: 25 °C for 10 min (primer annealing), 42 °C for 50 min (extension) and 70 °C for 15 min (inactivation of enzyme). The cDNA was subsequently diluted 1 in 30 times by adding 5 μ L of cDNA to 145 μ L nuclease-free water to a total volume of 150 μ L. Next, to perform quantitative-PCR, 5 μ L diluted cDNA is added with 0.6 μ L 5 μ M forward primer, 0.6 μ L 5 μ M reverse primer, 1.3 μ L nuclease-free water and 7.5 μ L 2X SensiMix SYBR Low-ROX kit (QT625-05). qPCR cycles are performed using qPCR machine (Mx3005P, Agilent) as follows: 95 °C for 10 min followed by 40 cycles of 95 °C for 30 s; 60 °C for 30 s; 72 °C for 30 s. Primers used for RT-PCR are listed in Table 2.3. Expression fold change is calculated as follows: $2^{-\Delta\Delta CT}$, where $\Delta\Delta CT = \Delta\text{knock-out}_{(\text{target gene- GAPDH})} - \Delta\text{empty}_{(\text{target gene- GAPDH})}$.

2.13 Optimization of qPCR primers

RT-PCR primers were designed using PrimerBank (<https://pga.mgh.harvard.edu/primerbank/>). ChIP-qPCR primers were designed using Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). All qPCR primers are tested for specificity prior to use for RT-qPCR or ChIP-qPCR. The presence of a single peak in the dissociation curve indicates that only one PCR product is present. More than one peak in the dissociation curve is likely due to more than one target has been amplified. Primers with more than one peak are discarded and redesigned. Only primers that give rise to one PCR product were used.

2.14 Optimization for ChIP-seq

For ChIP-seq purposes, DNA fragments are required to be 100-700 base-pair (bp). This is dependent on various factors, including the cell line, cell number and concentration, sonicator equipment model, cross-linker(s) used and the duration of cross-linking. Double cross-linking whereby a protein-protein linker such as disuccinimidyl glutarate (DSG) or ethylene glycol bis(succinimidyl succinate) (EGS) are used in addition to formaldehyde (FA), which links protein-DNA, typically requires longer cycles. We fixed the cell number per condition to be 20×10^6 cells using 1% FA (ThermoFisher, 28908).

To determine the number of sonication cycles required to obtain 100-700 bp DNA fragments for MOLM-13, MV4-11 and OCI-AML3 cross-linked with the above conditions for 5 and 10 min, we

assayed various cycles ranging from 5-25 cycles. For collection at each sonication cycle, 50 μ L of samples were extracted from the 2 mL suspension.

For 5 min cross-linked MOLM-13, the desired DNA-fragment is achieved between 8-14 sonication cycles (Figure 2.3A) and 8-17 cycles for 10 min cross-linked (Figure 2.3B). Therefore, we chose to use 10 cycles for both 5 min and 10 min cross-linked MOLM-13, under the condition of 20×10^6 cells per 2mL and cross-linking with 1% FA.

MV4-11 cross-linked for 5 min requires 8-11 cycles while 10 min require 11 cycles (Figure 2.4A). Note that, not enough DNA was loaded to the lane 8 (8 cycles, 10 min) and lane 10 (14 cycles, 10 min). Given the data available, we chose to use 10 cycles and 11 cycles for 5 min and 10 min cross-linked MV4-11, respectively.

OCI-AML3 cross-linked for 5 min requires 8-14 cycles while 10 min require 11-14 cycles (Figure 2.4B). Lane 4 (11 cycles, 5 min) and lane 8 (8 cycles, 10 min) were not loaded with sufficient DNA. However, we are still able to identify the optimal cycles. We opted for 11 cycles for both 5 min and 10 min cross-linked OCI-AML3.

Summary of optimal cycles for MOLM-13, MV4-11 and OCI-AML3 for 5 min and 10 min cross-linked with 1% FA, 20×10^6 cells in each condition:

| | 5 minutes | 10 minutes |
|----------|-----------|------------|
| MOLM-13 | 10 | 10 |
| MV4-11 | 10 | 11 |
| OCI-AML3 | 11 | 11 |

Additionally, we used the bioanalyzer (5067-4626, Agilent) to analysed the DNA fragment length of sonicated samples for 10 min cross-linked input samples of MOLM-13 and OCI-AML3. Although the sonication was not entirely efficient, as seen with a peak between 500-1000 bp in both cell lines, the majority of the DNA is within the target range (Figure 2.5).

To reverse cross-link the protein-DNA linkage and retain the DNA, the following was added to 10 μ L of samples and incubated at 65 °C for 2 h

| | |
|--------------------------------|------------|
| Nuclease free H ₂ O | 74 μ L |
| 5 M NaCl | 4 μ L |
| 20 mg/mL Proteinase K | 8 μ L |
| 10 mg/mL RNase A | 4 μ L |
| | <hr/> |
| | 90 μ L |

De-crosslinked DNA was then purified by QIAquick PCR purification Kit (Qiagen, 28104). 10 μ L purified DNA was added with 2 μ L 10X gel-loading buffer (Invitrogen, 10816015) and loaded into wells of 1.5%-2% agarose gel with a 100 bp DNA ladder (N3231, NEB).

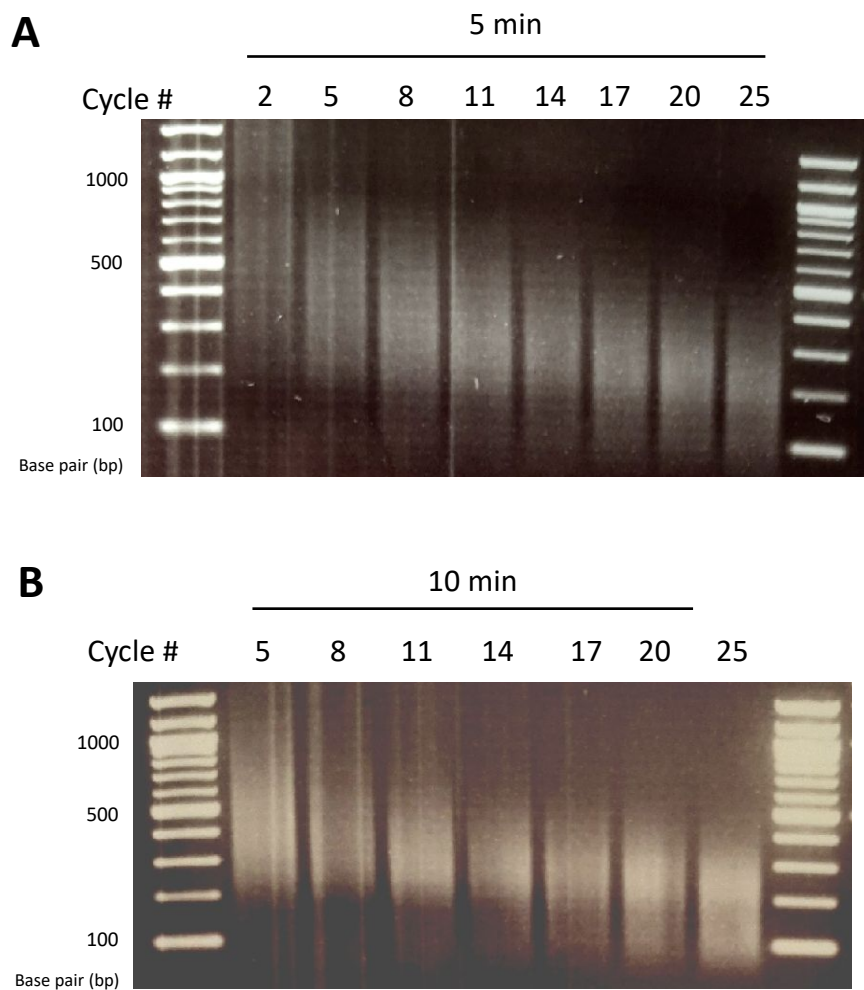


Figure 2.3 DNA fragments after sonication of cross-linked MOLM-13

MOLM-13 cells crosslinked with 1% formaldehyde for A) 5 min or B) 10 min and sonicated for various cycles to determine the number of cycles required to achieve the optimal 100-700 bp DNA fragments.

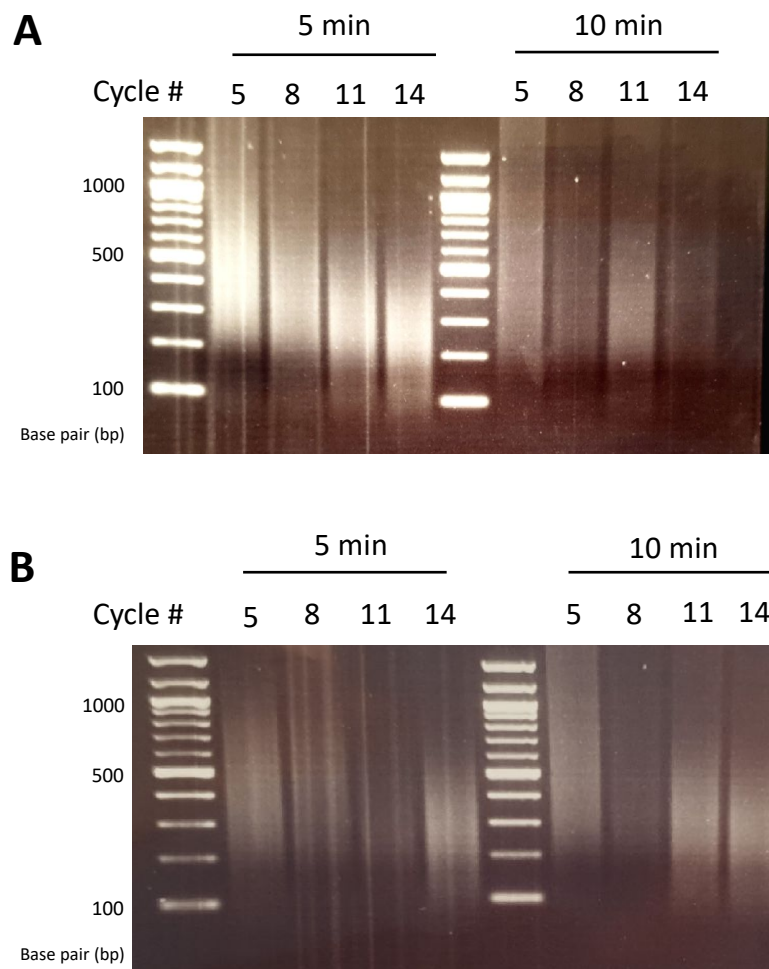


Figure 2.4 DNA fragments after sonication of cross-linked MV4-11 and OCI-AML3

A) MV4-11 and B) OCI-AML3 crosslinked with 1% formaldehyde for 5 min or 10 min and sonicated for various cycles to determine the number of cycles required to achieve the optimal 100-700 bp DNA fragments.

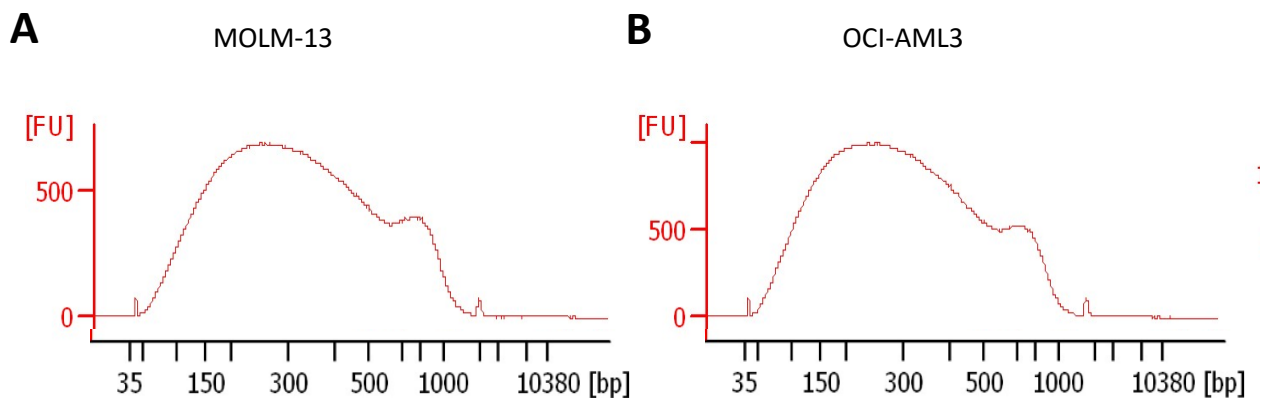


Figure 2.5 Analysis of input DNA fragments after the sonication of cross-linked MOLM-13 and OCI-AML3 by bioanalyzer

Cells were cross-linked with 1% formaldehyde for 10 min and sonicated for 10 and 11 cycles for MOLM-13 and OCI-AML3, respectively. Two distinct peaks are evident in both cell lines— one between 35-500 bp and the other between 500-1000 bp. Majority of the DNA fragments are within the 35-500 bp. These samples were submitted as input for ChIP-seq.

2.15 Validation of antibodies used in ChIP-seq

The success of ChIP-seq is dependent on the ability of the selected antibody to immunoprecipitate the target protein. We analysed two different KAT7 antibodies—monoclonal (58418, Cell Signalling) and polyclonal (ab70183, Abcam) for ChIP, followed by western blot analysis of the lysate. In addition, we also tested monoclonal and polyclonal H3K14ac antibodies (C15210005, C15410310 respectively, Diagenode) and H4K12ac antibody (07-595, EMD Millipore). Briefly, ChIP (see below) was performed until the washing step, after which, the supernatant was removed by centrifugation and the magnetic beads were retained using a magnetic rack. The beads (along with saved input) were then resuspended in SDS sample buffer, reducing agent as with standard western blot described above, and heated at 95 °C for 5 min followed by shaking at room temperature for 10 min. Samples were centrifuged to collect the liquid to the bottom of the Eppendorf tube and the magnetic beads were removed from the sample using a magnetic rack, the supernatant is transferred into a new Eppendorf tube. Samples were stored in -80 °C until the loading on to the gel.

A band from the ChIP with polyclonal KAT7 antibody (pAb KAT7), indicated by the yellow arrow, corresponds to the expected molecular weight of KAT7, reflecting successful immunoprecipitation of KAT7 protein (Figure 2.6). No band was observed for input, IgG control, immunoprecipitation with pAb H4K12ac antibodies, mAb and pAb H3K14ac antibodies (Figure 2.6). Insufficient loading could explain the lack of bands observed in the input lane. Absence of positive immunoprecipitation using the antibodies targeting the histone marks may be due to very little samples loaded, low quality of the antibodies or low abundance of the histone marks.

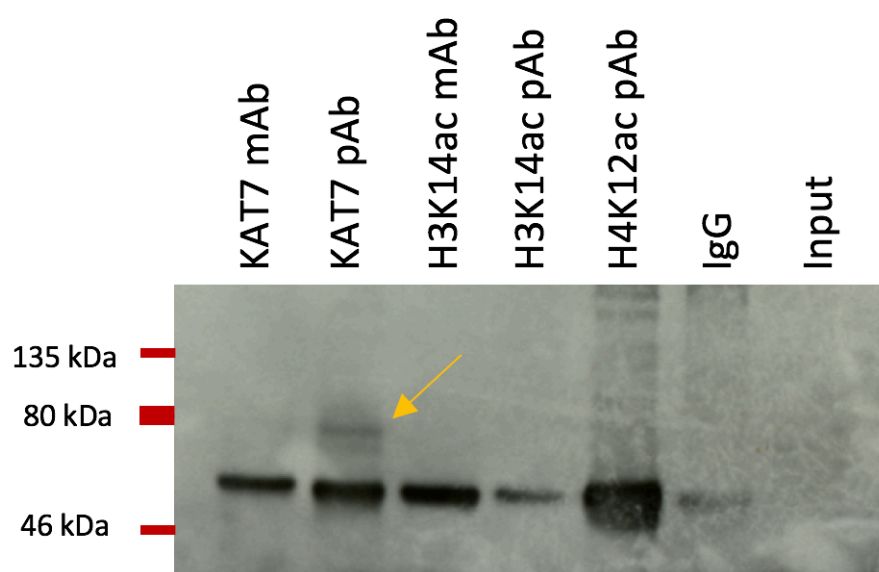


Figure 2.6 Immunoblot of ChIP using KAT7, H3K14ac and H4K12ac antibodies in MOLM-13

Chromatin immunoprecipitation using monoclonal (mAb) and polyclonal (pAb) antibodies of KAT7 and IgG (5 μ g antibody per ChIP) and other KAT7-mediated histone lysine acetylation targets (2 μ g antibody per ChIP). 20 \times 10⁶ MOLM-13 cells were cross-linked with 1% formaldehyde for 10 min and 5 min for KAT7 and histone ChIP, respectively. The bands at approximately 60 kDa are signals from the heavy chain of the antibodies used in the ChIP. The arrow indicates the successful pull-down of KAT7 protein using polyclonal KAT7 antibody.

2.16 ChIP-Seq sample preparation

AML cells were cross-linked in 1% FA at room temperature for 5 min (histones ChIP) or 10 min (non-histone ChIP) and subsequently quenched with 0.125 M Glycine (Sigma) at room temperature for 5 min. Cross-linked cells were washed twice with ice-cold PBS before sequential lysis with LB1 (50

mM Hepes, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100), LB2 (10 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), LB3 (10 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine) on ice for 10 min each. Lysed samples were sonicated for 10 cycles (MOLM-13) or 11 cycles (MV4-11 and OCI-AML3) using Bioruptor Pico instrument (Diagenode) for 10 min cross-linked cells and 10 cycles (MOLM-13, MV4-11) for 5 min cross-linked cells. Each cycle is 30 s ON/ 30 s OFF. Triton-X (Sigma) was added to the sonicated samples to a final concentration of 1% before centrifugation at 20000x g at 4 °C for 10 min. 10% of samples were kept as input. Lysates were incubated with antibody (Table 2.5) at 4 °C overnight, with rotation. Dynabeads Protein A/G (ThermoFisher) was added the next day and incubated at 4 °C for 4h, with rotation. ChIP samples were washed with RIPA wash buffer (50 mM HEPES, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-Deoxycholate) 3 times followed by wash with Annealing buffer (TE+ 50 mM NaCl). Elution Buffer (1% SDS, 50 mM Tris-HCl, 10 mM EDTA) was added to each ChIP samples and heated at 65 °C for 30 min. Beads were subsequently removed and samples were heated, at 14000 rpm, 65 °C overnight to reverse cross-linking whilst supplemented with 0.2 mg/mL RNase A (ThermoFisher, 10753721). 0.2 mg/mL Proteinase K (Life Technologies, 25530049) was added the next day to digest proteins by shaking at 450 rpm at 65 °C for 4h. DNA was purified by the QIAquick PCR purification Kit (Qiagen, 28104). PBS and LB1, LB2, LB3 and RIPA wash buffers were all supplemented with Sodium Butyrate (Sigma, 303410), cOmplete EDTA-free protease inhibitor cocktail (Roche, 4693132001) and PMSF (Sigma, 93482) immediately before use. Sequencing was performed on Illumina HiSeq v4 platform with 75-bp paired-end sequencing.

2.17 ChIP-qPCR

ChIP is performed as illustrated above with ChIP-seq. Input DNA is diluted 1 in 20 or 1 in 40, to obtain a concentration of <20 ng per 5 µL. KAPA SYBR Fast qPCR kit (KK4620) was used following the manufacturer's protocol. Thermal profile on qPCR machine (Mx3005P, Agilent) are as follows: 1 cycle of 95 °C for 3 min (enzyme activation); 40 cycles of 95 °C for 3 s (denaturing) and 60 °C for 20 s and 72 °C for 20 s (anneal and extend); 1 cycle of 95 °C for 1 min, 55 °C for 30 s, and 95 °C for 30 s (dissociation). Percentage Input is calculated as $100 \times 2^{\Delta Ct}$, where $\Delta Ct = Ct_{\text{Normalized input}} - Ct_{\text{ChIP}}$; $Ct_{\text{Normalized input}} = Ct_{\text{Input}} - \log_2(\text{input dilution factor})$; input dilution factor = fraction of input saved relative to each IP x dilution of input for qPCR. Sequences of primers used in ChIP-qPCR are listed in Table 2.4.

2.18 ChIP-seq processing

Reads of ChIP-seq sample and input DNA were mapped to the human genome GRCh38 using BWA-MEM version 0.7.17 (<https://github.com/lh3/bwa>) using default parameters. Duplicates were marked by Samtools mkdup (Li et al., 2009). Peaks (broad and narrow) were called by MACS2 version 1.4.1 (Zhang et al., 2008) using input DNA as control with parameters --broad-cutoff 0.01 for broad peaks, and -q 0.01 for narrow peaks. Broad and narrow peaks were merged into a union set. Broad peaks that overlap with one or more narrow peaks were removed. Locations of peaks (promoter, exon, intron or intergenic) were computed by customized scripts using Ensembl transcript annotation of GRCh38 version 91. Peaks were associated with promoter(s) if more than half of the peak length is located within ± 2 kb from the transcription start site. Promoter occupancies of a transcript were quantified as the highest MACS2 signal amongst all the peaks within the 2 kb window. If multiple isoforms exist, genic promoter occupancy was calculated as the highest signal amongst isoforms.

The public dataset used in this study: Geo Expression Omnibus GSM1845161 and GSM1845133.

2.19 RNA extraction and RNA-seq processing

RNA was extracted from AML cells with RNeasy Plus Mini Kit (Qiagen) according to the manufacturer's instructions. Sequencing was performed on Illumina HiSeq v4 platform with 75-bp paired-end sequencing.

RNA-seq reads were mapped to the human genome assembly GRCh38 using STAR version 2.5.0c (Dobin et al., 2013). tolerating mismatch rate of 0.01 and allowing maximal intron lengths of 50 kb. Read counts were calculated by STAR --quantMode using Ensembl annotation of GRCh38 version 91 and normalized as fragments per kilobase of gene length per million uniquely mapped reads (FPKM). Differential expression analysis was done by DESeq2 (Love et al., 2014) using paired sample design and significant genes were identified using adjusted *p-values* of 0.05 as the cut-off.

2.20 Data availability

The ChIP-seq and RNA-seq data have been deposited in the GEO database under accession numbers GSE133516.

2.21 Generation of Auxin Inducible Degron (AID) KAT7 protein degradation system and treatment with indole-3-acetic acid (IAA)

MOLM-13 and MV4-11 cells were introduced with pKLV-EF1aGFP2AKAT7_linkermAID-W by lentivirus transduction. Transduced AML cells were subsequently sorted for GFP-expression to enrich for population with KAT7-AID integration. KAT7-AID transgene is resistant to modification by gRNA (5) and gRNA (A10). KAT7-AID cells were subsequently transduced with lentivirus expressing pKLV2-U6gKAT7(A10)-PGKpuro2ABFP-W, as with gRNA-mediated knock-out described above, to knock-out endogenous *KAT7* and subsequently sorted for BFP-positive cells to enrich for knock-out population. GFP-BFP-double- positive cells were introduced with OsTIR1 F-box transgene via lentivirus transduction of pKLV-EF1a-mCherry-2A-OsTIR1-W. Cells with OsTIR1 transgene integration was enriched by sorting for the mCherry-positive cells. Indole-3-acetic acid (IAA) is the most common plant hormone of the auxin family. IAA (Sigma) was dissolved in water (W3500, Sigma) at a concentration of 500 mM and stored in -20 °C. MOLM-13 and MV4-11 AID cells were treated with 500 μ M IAA for 24h at 37 °C, 5% CO₂ for ChIP-qPCR experiments and 48 h at 37 °C, 5% CO₂ for differentiation assay.

2.23 IBET -151 treatment and viability assay

IBET-151 was dissolved in ethanol at a concentration of 10 mM. All cells tested was treated with 1 nM, 3 nM, 10 nM, 30 mM, 100 nM, 300 nM, 1 μ M and 3 μ M of IBET-151 (0.1% ethanol) for 3, 4 or 5 days, in 96-well plate. Viability of cells was assayed at each time point using CellTiter-Glow Luminescent Cell Viability Assay (7572, Promega), following the manufacturer's protocol. Briefly, both cells and reagent were equilibrated at room temperature for approximately 20 min. 100 μ L reagent was added to each well and mixed by pipetting. Next, the plates were shaken for 2 min and left at room temperature for 10 min. Luminescent was assayed using the Infinite F200 Pro plate reader.

2.24 Statistical analysis

Student's *t* tests (two-tailed) were used for statistical testing unless stated otherwise. Mean was calculated from at least three replicates, as indicated in each figure, and the error bars represent the standard deviation. Student's T-test (two-tailed) statistical testing was used unless indicated otherwise. *P* value of ≤ 0.05 were considered statistically significant.

Table 2.2 List of gRNAs used in this study

| gRNA | gRNA sequence (5'-3') |
|-------------|------------------------------|
| HIPK1_A1 | GAGACCTTAAGCCTCCACAG |
| HIPK1_A2 | GATATGTCTACAGACCTGGA |
| SIK3_A3 | GGATCCCACCGGACCTACAA |
| SIK3_A4 | GATCGACCGCACCATCGGCA |
| STK3_A5 | GGGTCTCTTAATATTGATAC |
| STK3_A6 | GACATTATTGCAGGATACAA |
| KAT6A_A7 | GCGGCGGATAATCACAAATC |
| KAT6A_A8 | GGAAC TAACGGTTCGAGTGA |
| KAT7_A9 | GGGAAAAAGTGGCTGAACTC |
| KAT7_A10 | GCCGCTATGAGCTTGATACC |
| KDM2A_A11 | GCCGAGTGGGGAATTTAAGC |
| KDM2A_A12 | GGCTCCTGACACAATCGGGG |
| KDM5B_B1 | GGCAGAATCTTACAACGAAT |
| KDM5B_B2 | GTCCAGGTATAGGTACACGC |
| KAT2A (1) | GGATGAGATAAACCGACTGC |
| SIK3 (1) | GATCTCGTAGTAGCCGATAC |
| SIK3 (2) | GGATCTCGTAGTAGCCGATA |
| SIK3 (3) | GACGAGATCGACCGCACCAT |
| SIK3 (4) | GTCGTAGTAGCCGATACGGG |
| SIK3 (5) | GCGGCTGGTGAGCATAACCG |
| SIK3 (6) | GCTTGAGTACCGGTGGACCT |
| KAT7 (1) | GGTGACTCGAGCAGATCGTC |
| KAT7 (2) | GGGTGACTCGAGCAGATCGT |
| KAT7 (3) | GCTTTAACAGGACACCTTAC |
| KAT7 (4) | GATCGAAGCTGTCTCTCCGT |
| KAT7 (5) | GGCTACCTGCATAATTTTCA |

Table 2.3 Sequences of primers used for RT-qPCR

| Primer sequence (5'-3') | | |
|-------------------------|-------------------------|-------------------------|
| Gene | Forward | Reverse |
| HOXA11 | TGCCAAGTTGTACTTACTACGTC | GTTGGAGGAGTAGGAGTATGTCA |
| HOXA10 | CTCGCCCATAGACCTGTGG | GTTCTGCGCGAAAGAGCAC |
| HOXA9 | TACGTGGACTCGTTCCTGCT | CGTCGCCTTGGACTGGAAG |
| HOXA7 | TCGTATTATGTGAACGCGCTT | CAAGAAGTCGGCTCGGCATT |
| HOXA5 | AACTCATTTTGC GGTCGCTAT | TCCCTGAATTGCTCGCTCAC |
| MEIS1 | GGGCATGGATGGAGTAGGC | GGGTACTGATGCGAGTGCAG |
| PBX3 | ATTACAGAGCCAAATTGACCCAG | TCTCGGAGAAGGTTTCATCACAT |
| GAPDH | GGAGCGAGATCCCTCCAAAAT | GGCTGTTGTCATACTTCTCATGG |
| BCL2 | GGTGGGGTCATGTGTGTGG | CGGTCAGGTACTCAGTCATCC |
| c-Myc | GGCTCCTGGCAAAAGGTCA | CTGCGTAGTTGTGCTGATGT |
| CDK6 | GCTGACCAGCAGTACGAATG | GCACACATCAAACAACCTGACC |
| JMJD1C | CAGGTCTCGTGCCAATCAAAA | GCTGTTGCTGGTGTGTATTCT |
| SEN6 | TCCTGTAAGGTAAAGTCGGCT | AGATAGAGGAGGAGTAGGCTGAT |

Table 2.4 Sequences of primers used for ChIP-qPCR

| Primer sequence (5'-3') | | |
|-------------------------|---------------------------|-------------------------|
| Gene | Forward | Reverse |
| control | GGCTCCTGTAACCAACCACTACC | CCTCTGGGCTGGCTTCATTC |
| SEN6 | TGCCTTTGTATAGGCCCGTC | ACGCCGCCTCCTTCTTAATC |
| PBX3 | CAGATGGGTCCGCCTTGTTTC | TGAACCGTGGGAAGTCAACA |
| PBX3 (intron) | AGTCAGAACCTCTCCGTGGT | CCTCTGAGAAACCGGAGTCG |
| MEIS1 | TTCCCCAAGTTAGCTGAGCG | AGGATCCGGTGGAGGAGAAA |
| MEIS1 (intron) | CTTCAGAGAACGATGCGGGT | AAGTCTAGCAACTGGGTGGC |
| JMJD1C | TGAAACAAAACCCAACGCGG | GGAGGGAGACGGAGCAGTA |
| CDK6 | TACTCTGGCGCTTTGTTGTG | CGCTGTAGGTAGCAGAGGT |
| BCL2 | GAGGAGGGCTCTTTCTTTCTTC | GCCTGTCCTCTTACTTCATTCTC |
| BCL2 (enhancer) | GAGCCCTCAACCTTGTGATAG | AAGGTAGCCCTGACCATAGA |
| HOXA9 | ATGCTTGTGGTTCTCCTC CAGTTG | CCGCCGCTCTCATTCTCAGC |
| HOXA10 | CGCAACCACCCAGCCAG | TTGTCCGCCGAGTCGTAGAGG |

Table 2.5 List of antibodies used for ChIP and Western Blot

| Antibody | Company | Catalogue No. | Application | Type | Host |
|---------------------------------|-----------------|----------------------|--------------------|-------------|-------------|
| KAT7 | Abcam | ab70183 | ChIP | pAb | rabbit |
| MLL1 | Bethyl Lab | A300-086A | ChIP | pAb | rabbit |
| BRD4 | Bethyl Lab | A301-985A | ChIP | pAb | rabbit |
| AF9 | Bethyl Lab | A300-597A | ChIP | pAb | rabbit |
| AF4 | Abcam | ab31812 | ChIP | pAb | rabbit |
| RNA Pol II S5 | Abcam | ab5131 | ChIP | pAb | rabbit |
| H4K5ac | Abcam | ab51997 | ChIP | mAb | rabbit |
| H4K8ac | Abcam | ab15823 | ChIP | pAb | rabbit |
| H3K14ac | Diagenode | C15210005 | ChIP | mAb | rabbit |
| IgG | Cell Signalling | 3900S | ChIP | mAb | rabbit |
| KAT7 | Cell Signalling | 58418 | WB | mAb | rabbit |
| H3K14ac | Diagenode | C15410310 | WB | pAb | rabbit |
| H3K23ac | Abcam | ab61234 | WB | pAb | rabbit |
| H4K5ac | Diagenode | C15410025 | WB | pAb | rabbit |
| H4K8ac | Diagenode | C15410103 | WB | pAb | rabbit |
| H3K12ac | Abcam | ab177793 | WB | mAb | rabbit |
| H4K16ac | Abcam | ab109463 | WB | mAb | rabbit |
| H3K9ac | Cell Signalling | 9649 | WB | mAb | rabbit |
| H3K4me3 | Diagenode | C15410003 | WB | pAb | rabbit |
| H3K27ac | Diagenode | C15410196 | WB | pAb | rabbit |
| H3K79me2 | Abcam | ab3594 | WB | pAb | rabbit |
| H3K27me3 | Diagenode | C15410195 | WB | pAb | rabbit |
| Actin | Sigma-Aldrich | A2228 | WB | mAb | mouse |
| Anti-rabbit IgG, HRP- linked | Cell Signalling | 7074 | WB | | goat |
| Anti-mouse IgG, HRP-linked | R&D Systems | HAF007 | WB | pAb | goat |

Chapter 3: Phenotypic validation of CRISPR-Cas9 AML-specific cell essential genes identify KAT7 as a potential novel therapeutic target for MLL-fusion AML

3.1 Introduction

Functional genomic screens using CRISPR-Cas9 is a powerful tool for identifying a wealth of novel drug targets for diseases such as oncology due to the ease of generating targeted bi-allelic mutations (Chen et al., 2015; Hart et al., 2015; Sánchez-Rivera and Jacks, 2015; Shi et al., 2015). Large-scale screenings will often identify false-positive hits (Sharma and Petsalaki, 2018), therefore major efforts are needed to distinguish biologically or therapeutically meaningful candidates from the extensive output of viability-based negative-selection screens.

Focusing on the dropouts generated from the AML CRISPR-Cas9 screens performed by Tzelepis et al., we aim to confirm and provide biological insights on promising genetic vulnerabilities in the hope to advance the current therapeutic landscape of AML. In this chapter, we perform downstream validation of a selection of genes which protein product encodes a catalytic activity, namely kinases and epigenetic modifiers, that have not previously been mechanistically studied from a therapeutic angle in the context of AML. In addition to proliferation which was the phenotypic readout of the AML CRISPR-Cas9 screens, we also assayed the effects of candidate gene knock-out on cellular and biological processes such as apoptosis, differentiation and cell cycle.

3.2 Assessing the Cas9 activity of human AML cell lines

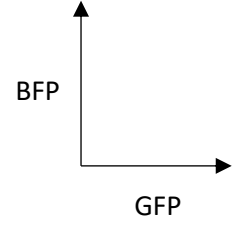
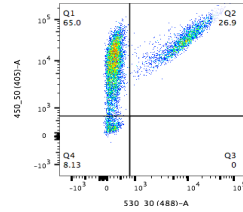
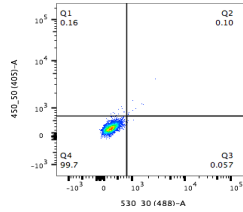
Prior to the validation of candidate genes from the AML CRISPR-Cas9 screens (Tzelepis et al., 2016), we first assessed the Cas9 activity of the AML cell lines that were used in the dropout screens (MOLM-13, MV4-11, OCI-AML2, HL-60 and OCI-AML3) as well as additional leukemic cell lines (THP-1, Nomo-1 and K562). The reporter construct *pKLV2-U6gRNA (target GFP)-PGK-BFP-T2A-GFP* (Tzelepis et al., 2016, refer to Chapter 2.6 for details), which expresses the *BFP* gene, *GFP* gene and gRNA targeting *GFP* (gGFP), was used. In the presence of active Cas9 proteins within the cell, the *GFP* gene sequence will be modified and consequently result in truncated or non-functional GFP protein, thereby lacking GFP expression. The proportion of active Cas9 cells within the population can, therefore, be assessed by calculating the proportion of BFP-positive population against the total

of all transduced cells (BFP-GFP-double-positive and BFP-single positive populations). Among the cell lines, K562 bulk population has the highest Cas9 activity (98.8%) and THP-1 has the lowest Cas9 activity (35.7%) (Figure 3.1). The Cas9 activity of other cell lines ranged from 70-80% in the bulk population— MOLM-13 (70.7%), MV4-11 (85.8%), OCI-AML2 (78.2%), OCI-AML3 (71.8%), HL-60 (87.2%) and Nomo-1 (82.5%) (Figure 3.1). The presence of Cas9-inactive cells within the cell population would predictively lead to lower gene editing efficiency, we therefore enriched for clones with high Cas9 activity by single-cell sub-cloning. After testing the Cas9 activity of individual clones, the clones with the highest Cas9 activity (>93%) were selected for use in validation and mechanistic studies (Figure 3.2). The importance of high Cas9 activity in functional studies is illustrated in Figure 3.3, using a previously validated gene, *KAT2A*. Loss of *KAT2A* was reported to induce CD11b expression, a marker for myeloid differentiation (Tzelepis et al., 2016) and we showed that lower Cas9 activity (pre-cloned population) results in no observable CD11b expression on day 4 post-transduction, while an increase in CD11b is seen in MOLM-13 with higher Cas9 activity (post-cloned population) (Figure 3.3).

The oncogenic mutations of the human leukemia cell lines used in this study are listed in Table 2.1.

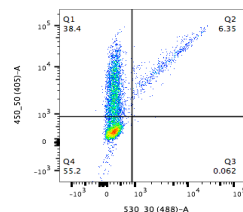
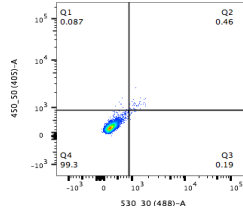
MOLM-13

70.7%



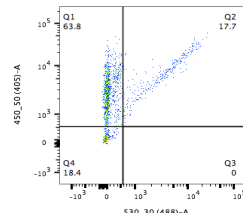
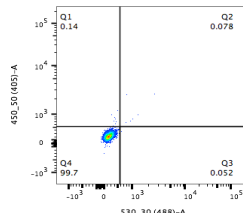
MV4-11

85.8%



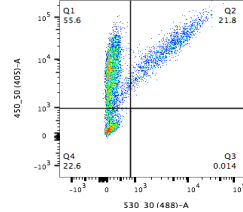
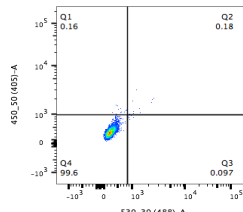
OCI-AML2

78.2%



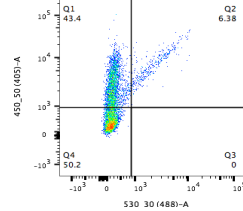
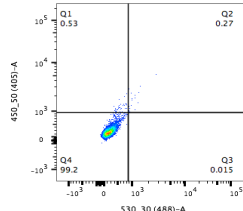
OCI-AML3

71.8%



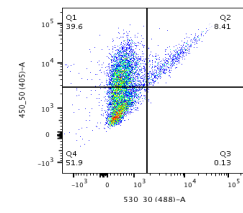
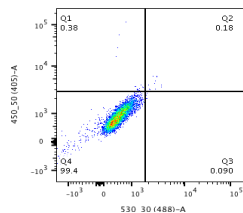
HL-60

87.2%



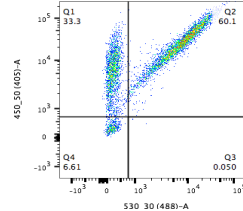
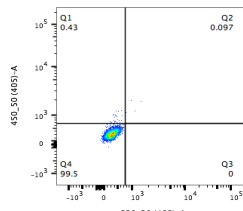
Nomo-1

82.5%



THP-1

35.7%



K562

98.8%

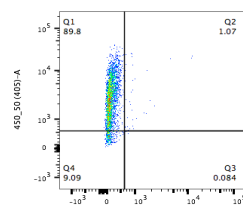
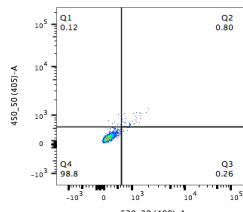


Figure 3.1 Cas9 activity of AML and CML cell lines before single-cell cloning.

The left panel shows mock (non-transduced, BFP-GFP double negative) population for gating purposes. The right panel shows reporter gRNA transduced population on day 3 post-transduction. The percentage of cells with active Cas9 activity is indicated and calculated by the proportion of BFP-positive population against the total of all transduced cells (BFP-GFP-double-positive and BFP-single positive population).

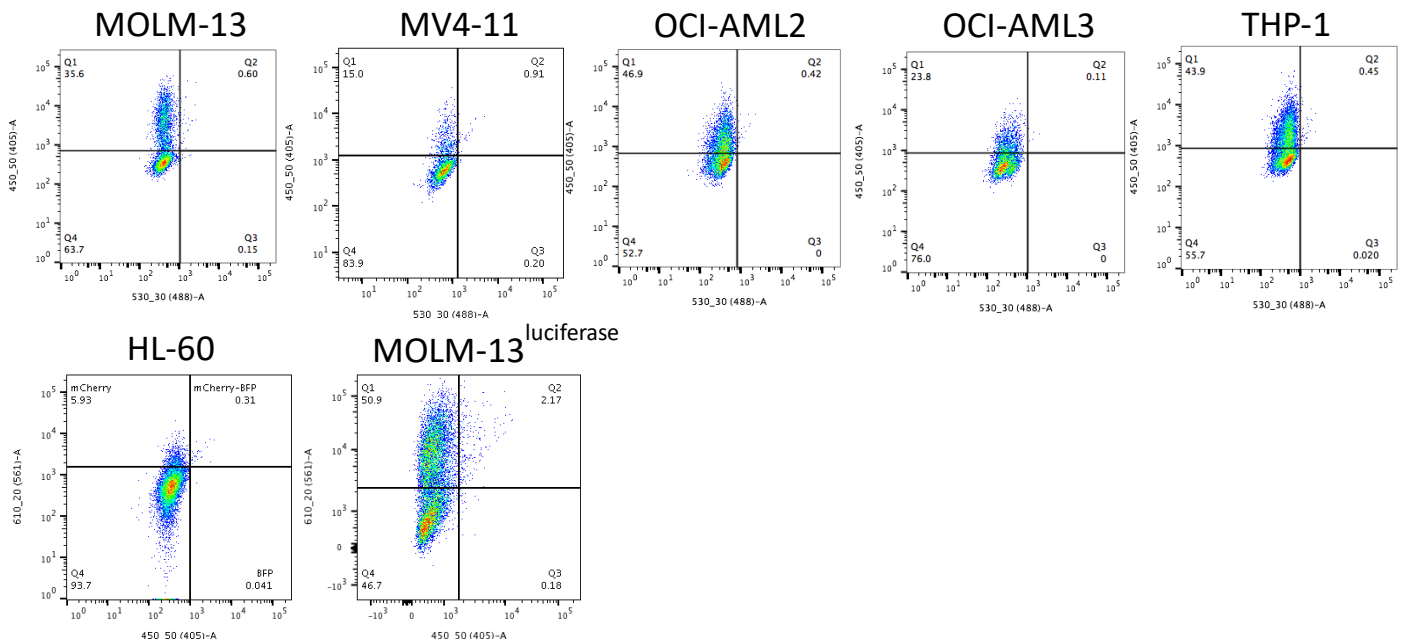


Figure 3.2 Cas9 activity of AML and CML cell lines after single-cell cloning.

MOLM-13, MV4-11, OCI-AML2, OCI-AML3 and THP-1 were transduced with lentivirus expressing the gGFP-BFP-GFP reporter construct and HL-60 and MOLM-13^{luciferase} were transduced with lentivirus expressing the gBFP-mCherry-BFP construct (refer to Chapter 2.6 and Figure 2.1). Colonies with >93% Cas9 activity were selected for downstream experiments.

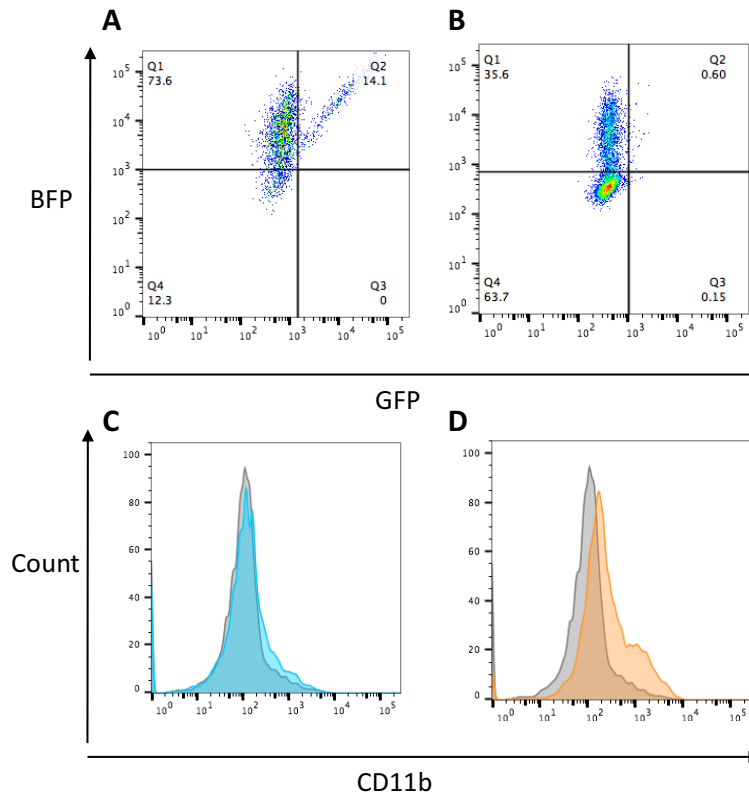


Figure 3.3 Cas9 activity and effect on functional phenotype

A-B) MOLM-13 cells transduced with Cas9 reporter virus reveal a difference in Cas9 activity in old (pre-cloned) and new clones (post-cloned). BFP and GFP double-positive reflects cells with inactive Cas9, whilst cells with BFP positive only are Cas9 active. A) 83.9% of cells have active Cas9 in old clone, B) 98.3% Cas9 activity in new clone. C-D) gRNA-mediated knock-out of KAT2A induce different levels of CD11b expression in the old (C, blue histogram) and the new (D, orange histogram) MOLM-13 clones, compared to empty control (grey histogram).

3.3 Validation of AML-specific druggable candidate genes from CRISPR-Cas9 dropout screens

In the AML CRISPR-Cas9 screens, 492 dropouts were AML-specific cell essential genes (not a dropout in HT-29 colon cancer line and HT-1080 fibrosarcoma line), 227 were “druggable” (Tzelepis et al., 2016). The Drug Gene Interaction Database (DGIdb) define druggable genes as genes/proteins that have potential or known to interact with drugs (Cotto et al., 2018). Within the “druggable” categories, many targets are either implicated in cancer and/or inhibitors are in clinical trials such as DOT1L, MCL1 and BCL2 (Caenepeel et al., 2018; Hird and Tron, 2019; Klaus et al., 2014; Li et al., 2019; Montero and Letai, 2018; Perini et al., 2018; Stein et al., 2018). We selected histone modifiers and kinases that have not been linked to AML or broadly in cancer for validation (Table 3.1). The

candidates include four kinases (*HIPK1*, *STK3*, *SIK3* and *RSK1*) and four epigenetic modifiers (*KDM2A*, *KDM5B*, *KAT6A* and *KAT7*). The 2 best performing gRNAs from the CRISPR-Cas9 screen performed by Tzelepis et al. (unpublished data) were used for validation in selected two cell lines, MOLM-13 and OCI-AML3. As all candidate genes, except for *SIK3* and *RSK1*, were essential in MOLM-13 but not in OCI-AML3 according to the screen results, these 2 cell lines could serve to validate essentiality and non-essentiality of the selected candidates.

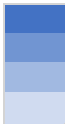
| | MOLM-13 | MV4-11 | HL-60 | OCI-AML2 | OCI-AML3 | HT-1080 | HT-29 | |
|--------------|---------|--------|-------|----------|----------|---------|-------|---|
| HIPK1 | 0.08 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | FDR  0 0.2 |
| STK3 | 0.15 | 1.00 | 1.00 | 0.87 | 1.00 | 0.95 | 1.00 | |
| SIK3 | 0.05 | 0.70 | 1.00 | 1.00 | 0.16 | 1.00 | 1.00 | |
| RSK1 | 0.01 | 1.00 | 0.02 | 0.00 | 0.00 | 1.00 | 1.00 | |
| KDM2A | 0.09 | 0.65 | 0.60 | 0.26 | 0.69 | 1.00 | 1.00 | |
| KDM5B | 0.15 | 0.49 | 1.00 | 0.53 | 0.22 | 1.00 | 1.00 | |
| KAT6A | 0.15 | 0.07 | 0.78 | 0.04 | 0.42 | 1.00 | 1.00 | |
| KAT7 | 0.05 | 0.16 | 1.00 | 0.05 | 0.48 | 0.53 | 0.73 | |

Table 3.1 List of novel histone modifiers and kinases that are AML-specific cell-essential genes.

Values indicated are the false discovery rate (FDR) of each gene in the respective cell lines (data derived from Tzelepis et al.). FDR < 0.2 is considered significant cell-essential genes (highlighted in blue). Homeodomain interacting protein kinase 1 (*HIPK1*); Serine/threonine-protein kinase 3 (*STK3*); Salt inducible kinase 3 (*SIK3*); Ribosomal protein S6 kinase (*RSK1*), also known as *RPS6KA1*; lysine demethylase 2A (*KDM2A*); lysine demethylase 5B (*KDM5B*); lysine acetyltransferase 6A (*KAT6A*); lysine acetyltransferase 7 (*KAT7*).

Knock-out of *HIPK1* and *STK3* showed a weak suppressive effect on the proliferation of MOLM-13 (Figure 3.4A, C) and had no effect in OCI-AML3 (Figure 3.4B, D). *KAT6A*, *SIK3* and *RSK1* knock-out confirmed the screen results. Targeting *SIK3* and *KAT6A* in MOLM-13 had a strong and moderate suppressive effect, respectively (Figure 3.4E, K) but weak and no effect on proliferation in OCI-AML3, respectively (Figure 3.4F, L). *RSK1* showed a slight effect in both cell lines (Figure 3.4G-H). *KDM2A* showed weak phenotype in MOLM-13 (Figure 3.4I) and OCI-AML3 (Figure 3.4J). *KDM5B* knock-out did not affect the proliferation of MOLM-13 or OCI-AML3 (Figure 3.4M-N). *KAT7* knock-out in OCI-AML3 showed a slight decrease in proliferation over a period of 12 days (Figure 3.4P); however, proliferation was strongly decreased in MOLM-13 (Figure 3.4O). In summary, both *SIK3* and *KAT7* knock-out showed a prominent decrease in proliferation and therefore selected these two genes for further investigations.

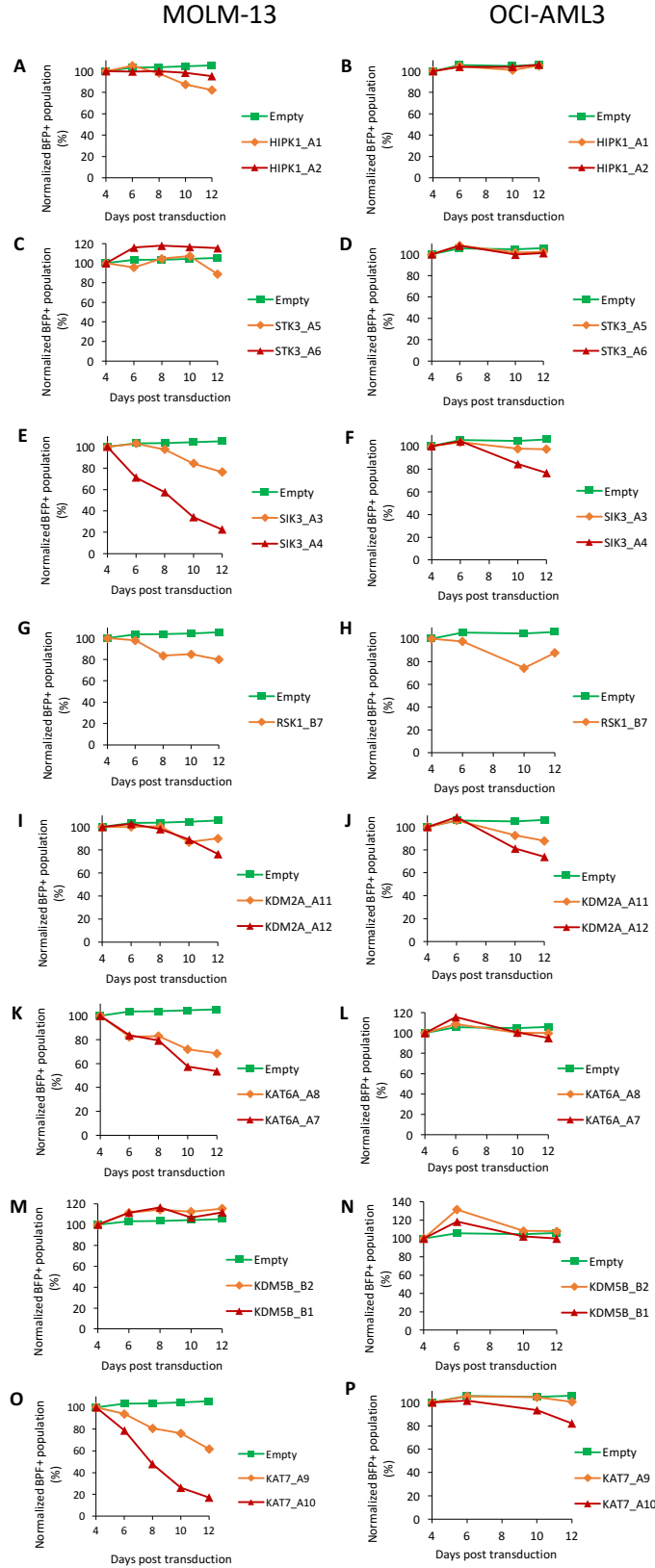


Figure 3.4 Validation of candidate hits from the CRISPR-Cas9 dropout screen in MOLM-13 and OCI-AML3

The left and right panel depicts MOLM-13 and OCI-AML3 knock-outs respectively. Competition-based proliferation assay between transduced (BFP-positive) and non-transduced (BFP-negative) cells. Two gRNAs for each gene were selected from the CRISPR library and gRNA was delivered to AML cells by lentivirus transduction. The proportion of BFP-positive population is normalized to day 4. Each data point is the average of triplicates.

3.4 Proliferation of KAT7 and SIK3 knock-out in various AML and CML cell lines

As *SIK3* and *KAT7* knock-out showed the most drastic effect on proliferation in MOLM-13, we further tested the effect of these two genes in 3 additional AML cell lines (MV4-11, OCI-AML2 and THP-1) and a chronic myeloid leukaemia (CML) cell line (K562) by using the best performing gRNA among the one selected from the library—gSIK3 (A4) and gKAT7 (A10) (Figure 3.5). *SIK3* knock-out showed a strong reduction in the proliferation in MOLM-13 and MV4-11 cells (Figure 3.5A, D), a moderate decrease in OCI-AML2 and THP-1 proliferation (Figure 3.5B, E) and no apparent effect in OCI-AML3 and K562 (Figure 3.5C, F). Knock-out of *KAT7* strongly reduced the proliferation of MOLM-13, MV4-11, OCI-ALM2 and THP-1 cells (Figure 3.5A, B, D, E), whilst minimal effect was observed in OCI-AML3 and K652 (Figure 3.5C, F). According to the dropout screens (Tzelepis et al., 2016), *SIK3* was not essential in MV4-11 or OCI-AML2 but was found to be essential in the validation experiments. *KAT7* knock-out was validated in all cell lines used in the screen and additionally found to be a vulnerability in THP-1. To address the potential off-target effects of the gRNAs chosen from the genome-wide library, we designed and tested additional gRNAs for *KAT7* (gKAT7) and *SIK3* (gSIK3), in MOLM-13 and MV4-11 (Figure 3.6). Among the gKAT7, only gKAT7 (5) has a comparable effect on proliferation as gKAT7 (A10) which was derived from the CRISPR library (Figure 3.6A). gSIK3 (A4) from the genome-wide library, gSIK3 (1), gSIK3 (2), gSIK3 (3), and gSIK3 (4) all have a similar effect on proliferation in MOLM-13 and MV4-11 (Figure 3.6B). For the purpose of further validation, we used gKAT7 (A10) and gSIK3 (A4) for downstream experiments. Furthermore, MOLM-13 was selected as the model cell line initially to investigate the cellular mechanism underlying the proliferation phenotype as it showed the strong anti-proliferative phenotype and the wild-type cells grow well in cell culture.

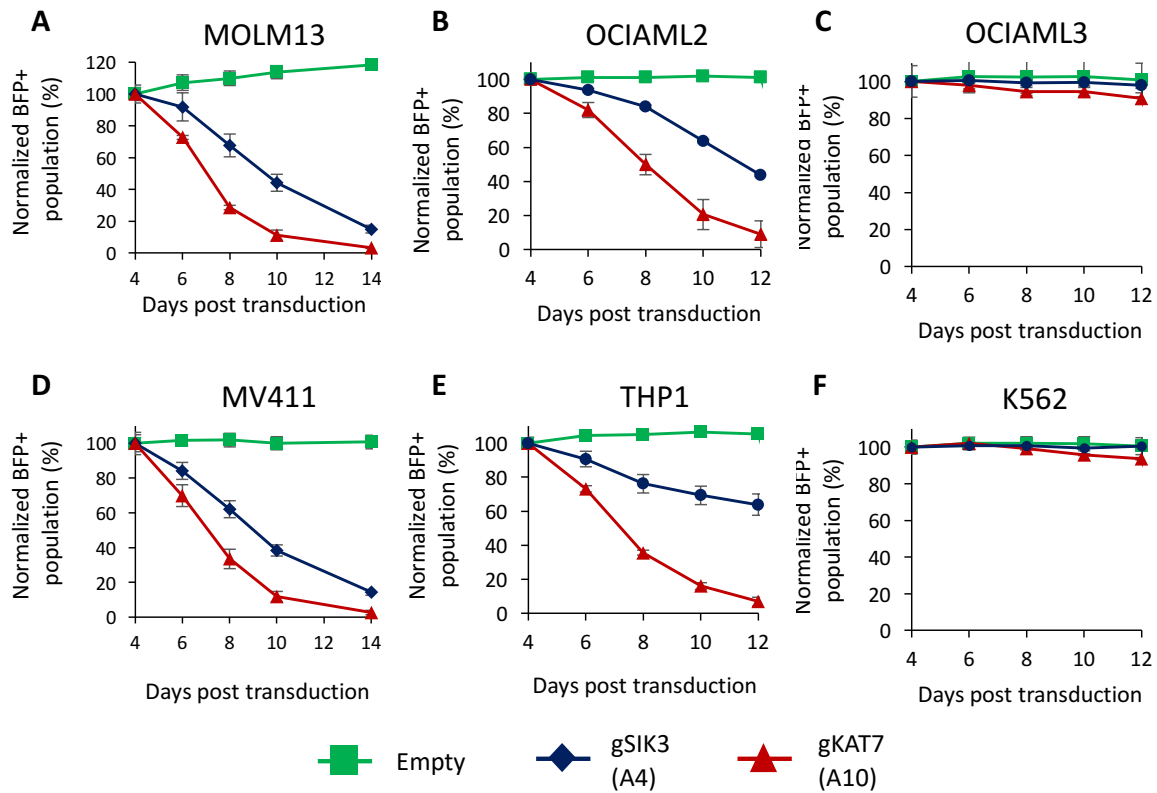


Figure 3.5 Proliferation of SIK3 and KAT7 knock-out in additional cell lines

Competition-based proliferation assay between transduced (BFP-positive) and non-transduced (BFP-negative) cells. Best performing gRNA for SIK3 and KAT7 was delivered by lentivirus. The proportion of BFP-positive population is normalized to day 4. Mean \pm standard deviation, $n=3$. KAT7 knock-out showed a strong decrease in proliferation in A) MOLM-13, B) OCI-AML2, D) MV4-11, and E) THP-1 whilst no effect on the proliferation of C) OCI-AML3 and F) K562. SIK3 knock-out has a strong reduction in the proliferative ability of A) MOLM-13 and D) MV4-11, moderate effect in B) OCI-AML2 and E) THP-1 and no effect in C) OCI-AML3 and F) K562.

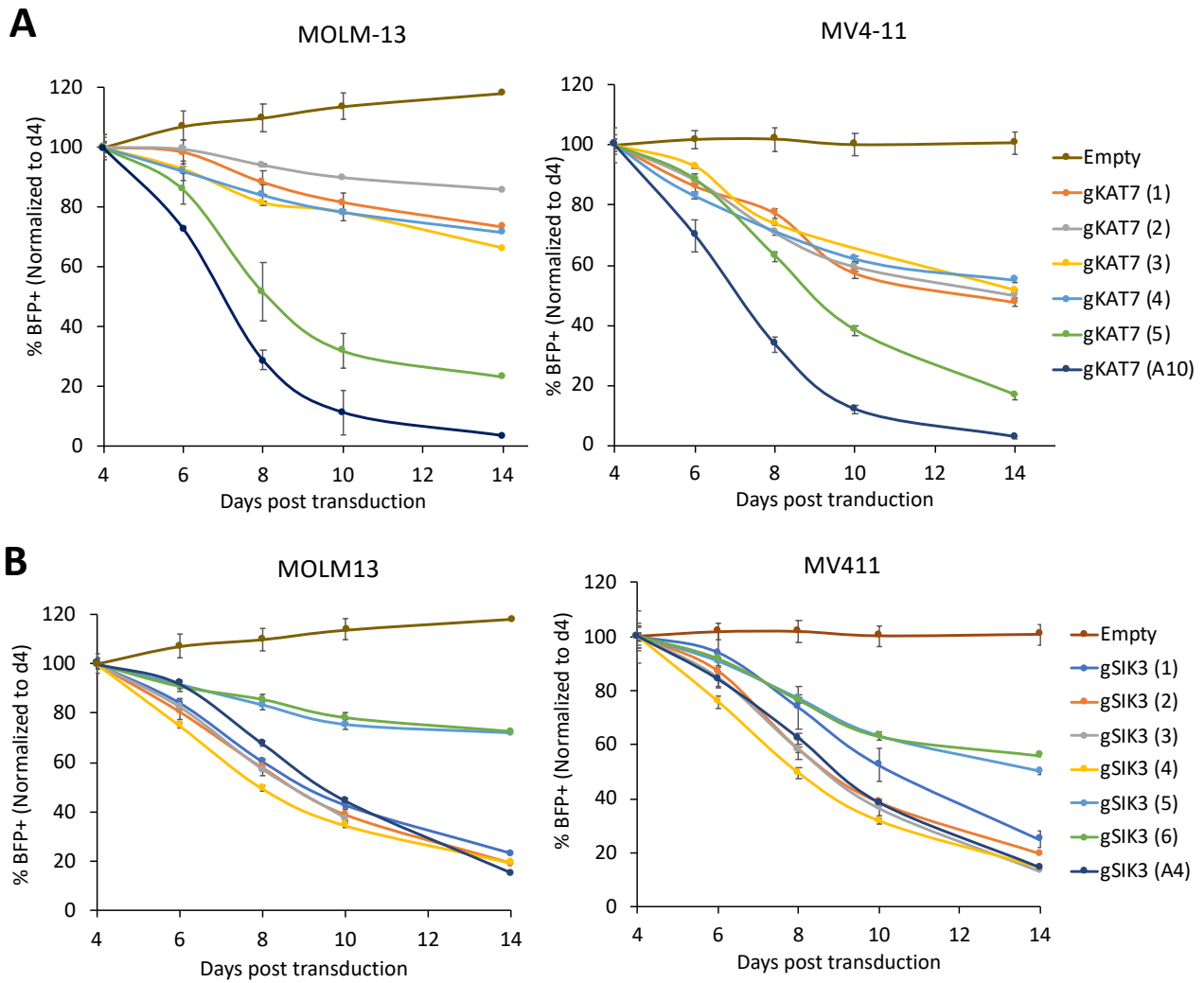


Figure 3.6 Proliferation of MOLM-13 and MV4-11 with different gRNAs targeting KAT7 (gKAT7) and SIK3 (gSIK3).

Competition-based proliferation assay between transduced (BFP-positive) and non-transduced (BFP-negative) cells. A) gKAT7 B) gSIK3. The proportion of BFP-positive population is normalized to day 4.

Mean \pm standard deviation, $n=3$.

3.5 KAT7 and SIK3 knock-out in MOLM-13

3.5.1 Apoptosis

To understand the cellular processes underlying the proliferation phenotype observed in *SIK3* and *KAT7* knock-out, we asked whether apoptotic cell death was responsible for the decrease in proliferation of the AML cells (Figure 3.7). Knock-out of *KAT7* in MOLM-13 results in a 13.7-times increase in Annexin V-positive cells compared to the empty negative control ($p=0.03$) on day 9 post-transduction (Figure 3.7A, C, E). This is four times more than the positive control, *KAT2A* knock-out (Figure 3.7A, D, E). On a smaller magnitude, *SIK3* knock-out in MOLM-13 has a 1.8-fold increase in Annexin V-positive cells (Figure 3.7A, B, E) but this difference is not statistically significant ($p=0.15$).

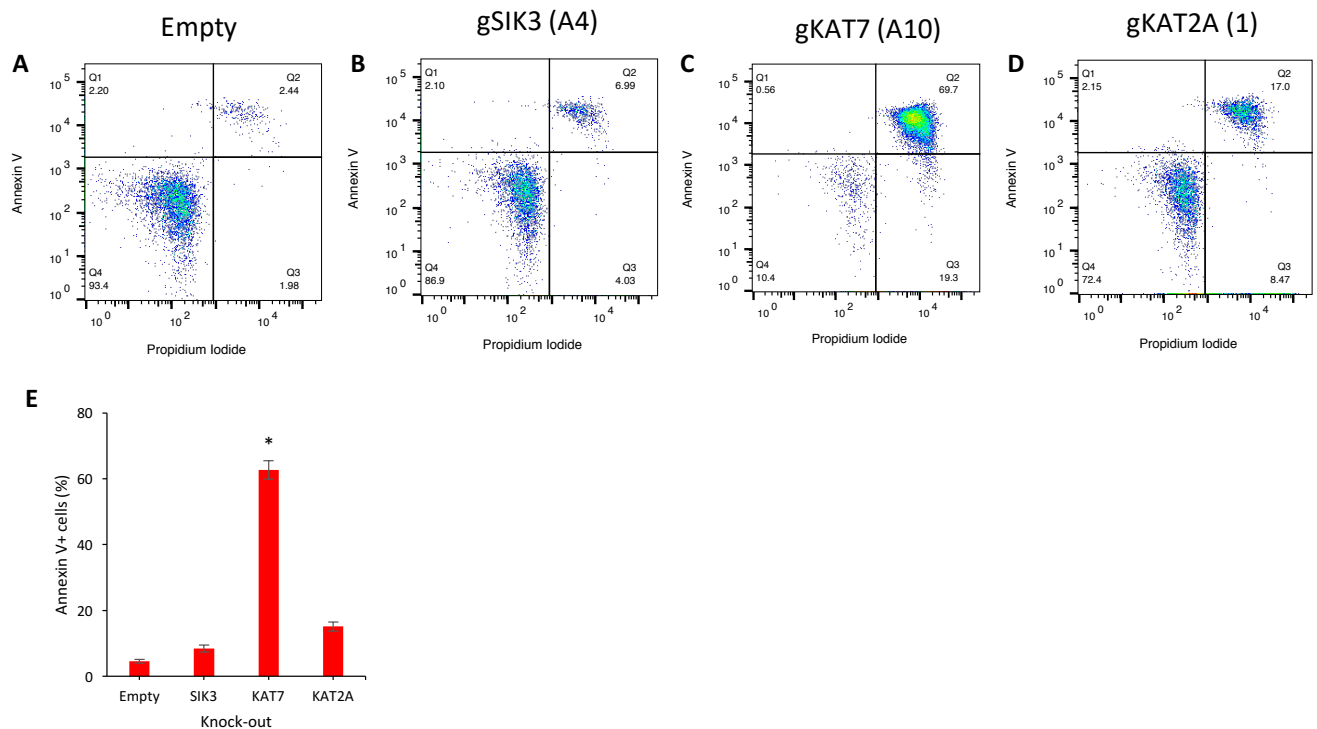


Figure 3.7 Apoptosis of *SIK3* and *KAT7* knock-out in AML cell lines

MOLM-13 was transduced by lentivirus expressing gSIK3 (A4), gKAT7 (A10), gKAT2A (1) or empty control. BFP-positive (knock-out) cells were sorted on day 3 and levels of apoptosis measured on day 9. A-D) MOLM-13 transduced with gRNA targeting A) empty, B) *SIK3*, C) *KAT7*, and D) *KAT2A* and stained for Annexin V and Propidium Iodide (PI). E) Significant increase in Annexin V-positive cell population in *KAT7* knock-out relative to empty day 7 post-transduction. Mean \pm standard deviation, $n=3$. Two-tailed t -test, *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$.

3.5.2 Differentiation

Previous studies found that the inhibition of *KAT2A* induces myeloid differentiation in AML (Tzelepis et al., 2016). We thus investigated whether knock-out of *SIK3* or *KAT7* induces differentiation in AML cells, by assessing the expression of two myeloid differentiation markers, CD11b and CD13, on day 5, 6 and 7 post-transduction. *KAT2A* knock-out was used as a positive control. Increase in CD11b expression, relative to the empty transduced cells, is detectable in *KAT7* knock-out and *KAT2A* knock-out but not in *SIK3* knock-out MOLM-13 cells (Figure 3.8). CD11b levels in knock-out also increase by approximately 10% each day from day 5 to day 7 post-transduction in *KAT7* knock-out cells (Figure 3.8). A marginal increase in CD13 expression, compared to empty control, can be seen in all knock-out cells by day 7 (Figure 3.9), however, not as striking as CD11b. Further repeats showed that genetic ablation of *KAT7* strongly induced CD11b expression in MOLM-13 and leads to 6 times more differentiated cells on day 7 (Figure 3.10B, G; $p=0.003$) than the empty control. This is 2 times higher than the positive control *KAT2A* knock-out (Figure 3.10C, G). CD13 expression increased by 2.6-fold in *KAT7* knock-out, compared to the empty control (Figure 3.10E, H), but this increase was slightly weaker than *KAT2A* knock-out (Figure 3.10F, H). *SIK3* knock-out did not show a drastic increase in CD11b or CD13 differentiation markers (Figure 3.10A, D). The levels of CD13 expression is comparable between the knock-outs (Figure 3.10H).

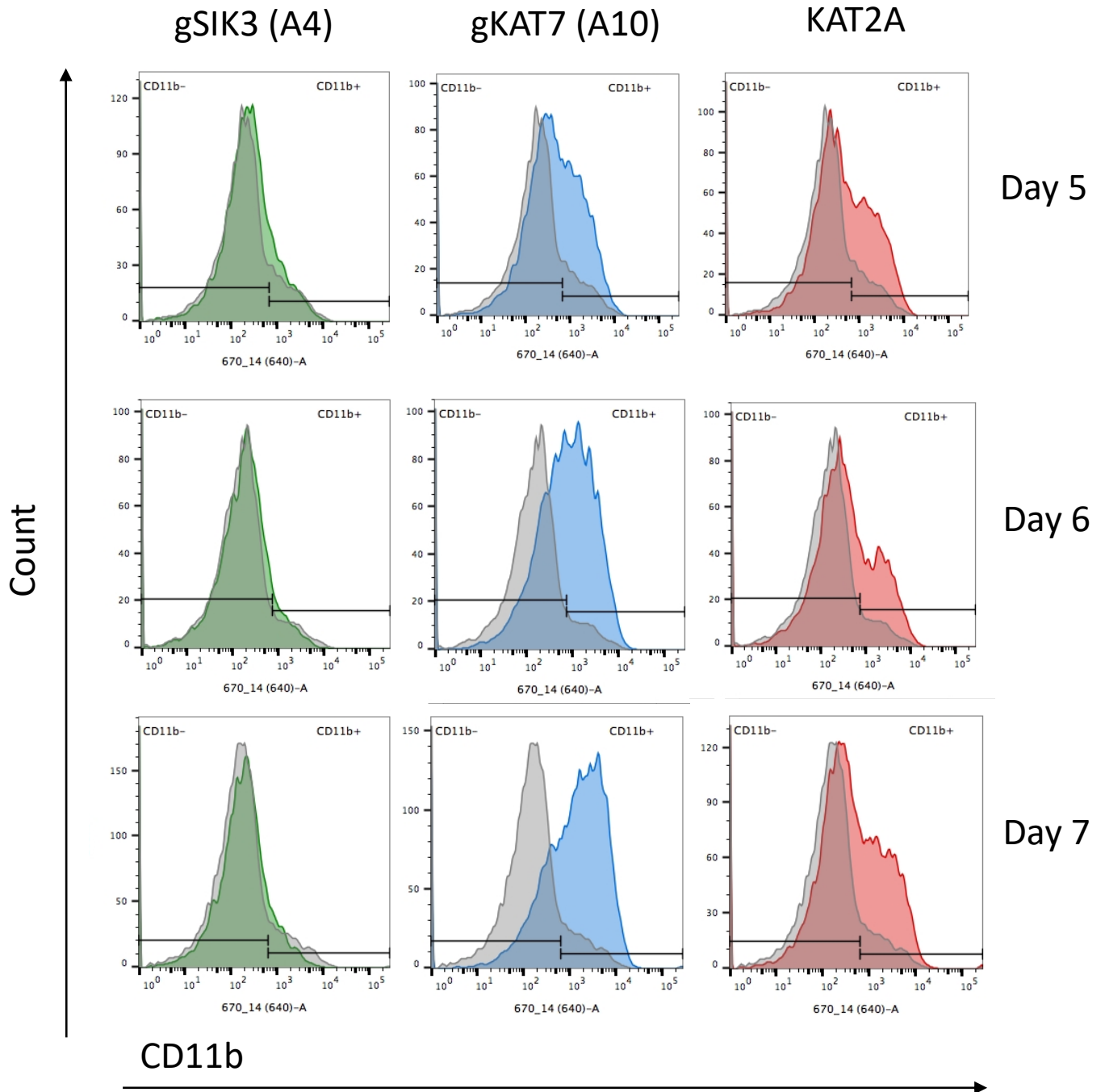


Figure 3.8 Time course CD11b differentiation assay in SIK3, KAT7 and KAT2A knock-out MOLM-13 cells

MOLM-13 was transduced by lentivirus expressing gSIK3 (A4), gKAT7 (A10), gKAT2A (1) or empty control. BFP-positive (knock-out) cells were sorted on day 3 post-transduction and assayed on day 5, 6 and 7. CD11b expression in SIK3 knock-out (green), KAT7 knock-out (blue) and KAT2A knock-out (red), against empty (grey).

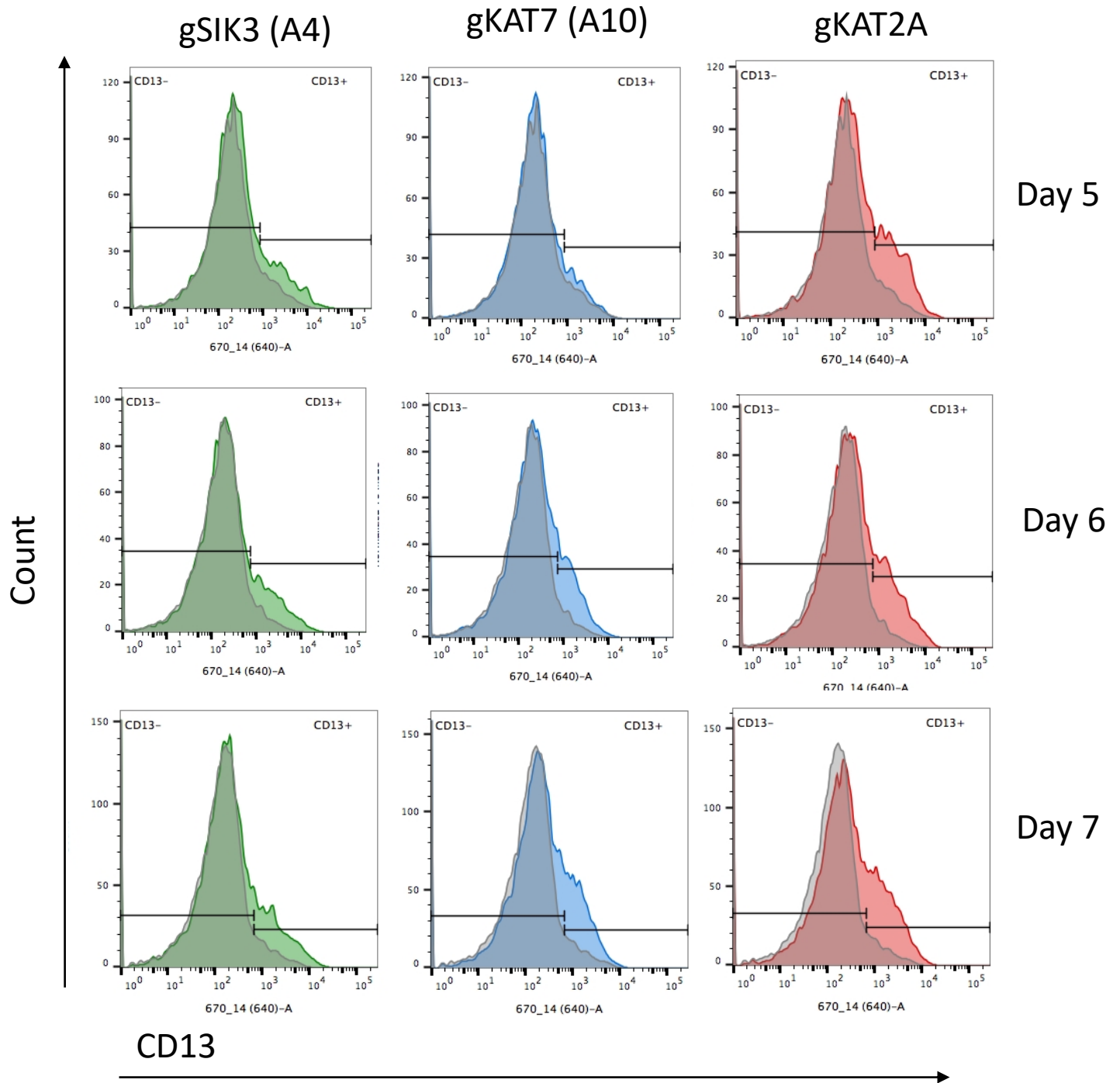


Figure 3.9 Time course CD13 differentiation assay in SIK3, KAT7 and KAT2A knock-out MOLM-13 cells

MOLM-13 was transduced by lentivirus expressing gSIK3 (A4), gKAT7 (A10), gKAT2A (1) or empty control. BFP-positive cells (knock-out) were sorted on day 3 post-transduction and assayed on day 5, 6 and 7. CD13 expression in SIK3 knock-out (green), KAT7 knock-out (blue) and KAT2A knock-out (red), against empty (grey).

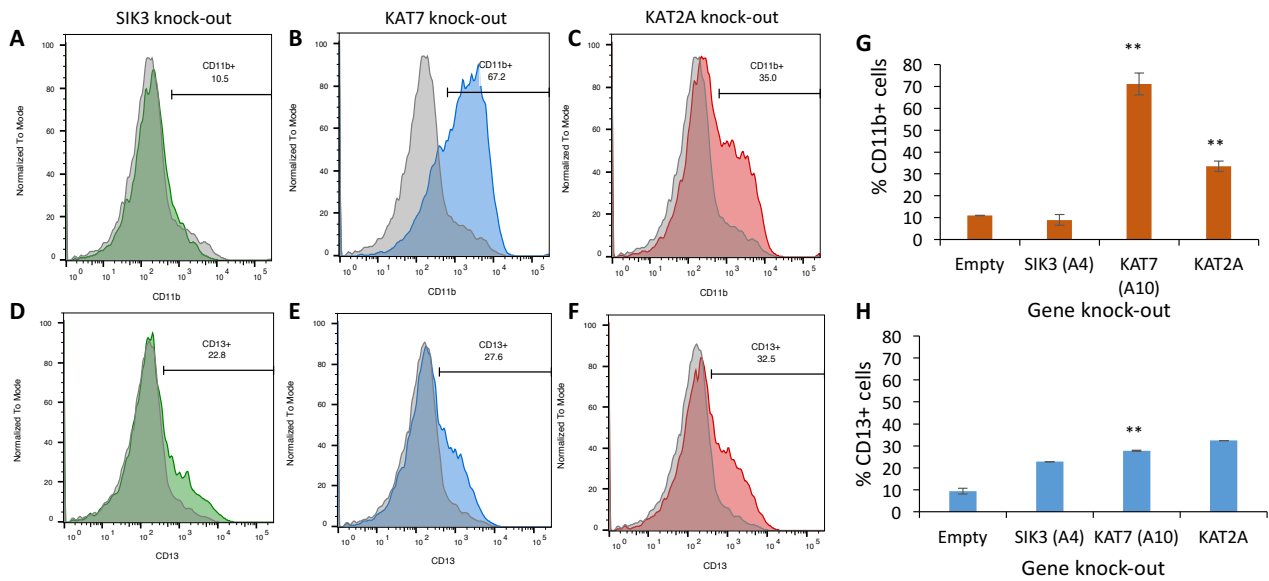


Figure 3.10 CD11b and CD13 cell surface marker expression of SIK3 and KAT7 knock-out in MOLM-13

MOLM-13 was transduced by lentivirus expressing gSIK3 (A4), gKAT7 (A10), gKAT2A (1) or empty control. BFP-positive (knock-out) cells were sorted on day 3 and assayed on day 7. Mean \pm standard deviation, $n=2$ (except KAT7-knock-out: CD11b, $n=5$; CD13, $n=3$). A-C) CD11b expression in knock-out (coloured), against empty (grey), with 10% spontaneous differentiation in wild-type cells. D-F) CD13 expression in knock-out (coloured), against empty (grey), with 10% spontaneous differentiation in wild-type cells. G) Relative CD11b levels in wild-type and knock-out, with KAT7 showing statistical significance. H) Relative CD13 levels in wild-type and knock-out. Two-tailed t -test, *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$.

3.5.3 Cell cycle

An alternative explanation for the observed proliferative phenotype could be due to cell cycle arrest. In particular, we saw a strong decrease in proliferation in MOLM-13 upon the knock-out of SIK3, but apoptosis assays revealed non-significant increase cell death. This prompted us to look at the cell cycle status of the knock-outs. Using propidium iodide (PI) to stain DNA, we detected an increase in the number of cells in Sub G1, S and G2/M phases and a corresponding decrease in G1 phase of MOLM-13 SIK3 knock-out compared to empty on day 8 post-transduction (Figure 3.11). However, this result is not consistently observed across independent biological replicates.

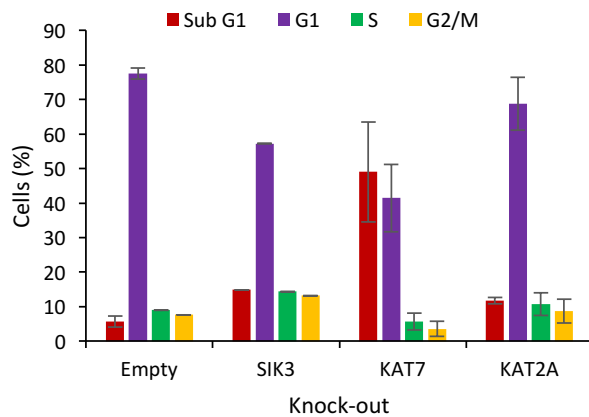


Figure 3.11 Cell cycle analysis of SIK3 and KAT7 knock-out in MOLM-13

MOLM-13 was transduced by lentivirus expressing gSIK3 (A4), gKAT7 (A10), gKAT2A (1) or empty control. BFP-positive (knock-out) cells were sorted on day 3 and cells are stained with PI on day 8.

Mean \pm standard deviation, $n=2$. MOLM-13 cells with either SIK3, KAT7 or KAT2A knock-out on day 8 post-transduction

3.6 KAT7 knock-out in additional AML cell lines

The data thus far showed gRNA-mediated KAT7 knock-out and SIK3 knock-out induces prominent proliferation defects in multiple AML cell lines. KAT7 knock-out, in addition, showed increased apoptosis and CD11b expression in MOLM-13. A slight increase in apoptosis was observed in MOLM-13 SIK3 knock-out whilst no changes in CD11b expression was observed. Due to these striking phenotypes as a result of genetic loss of KAT7, we thus chose to focus our efforts on understanding the mechanism of essentiality of KAT7 in AML. After selecting, designing and testing various gRNAs that target KAT7, we opted to use gKAT7 (A10) and gKAT7 (5) for further experiments as these two gRNAs resulted in the most rapid loss of proliferation across cells in which KAT7 is essential.

3.6.1 Apoptosis

Focusing on KAT7 as a vulnerability in AML, we investigated the effect of KAT7 knock-out in MV4-11, OCI-AML2, THP-1 and OCI-AML3, in addition to MOLM-13. Loss of KAT7 in MV4-11, OCI-

AML2 and THP-1 result in a significant increase in the proportion of Annexin V-positive cells on day 9 post-transduction (Figure 3.12). There is no significant increase in apoptotic cells in OCI-AML3 following the genetic ablation of *KAT7* (Figure 3.12). These data are consistent with the proliferation phenotype described above in Chapter 3.3. Notably, levels of annexin-V in THP-1 knock-out compared to empty control was not as drastic as MOLM-13, MV4-11 and OCI-AML2. This could be due to the fact that THP-1 possess the p53 mutation (Sugimoto et al., 1992). Wild-type p53 is a positive regulator of apoptosis, among many other cellular processes (Aubrey et al., 2018). In the case of THP-1, the presence of p53 mutant may compromise the protein's ability to induce maximal apoptosis under stress, such as coping with the loss of *KAT7*. This may suggest that *KAT7*-induced apoptosis is mediated through p53. Existing studies suggest that p53 negatively regulates *KAT7* in Saos-2 osteosarcoma cells and mouse embryonic fibroblast, most likely through the binding of p53 protein to the catalytic domain of *KAT7* (Iizuka et al., 2008). The possibility of *KAT7* in regulating p53 has been briefly explored in this study. Iizuka et al., compared the ability of *KAT7* to acetylate histone or p53 as substrates and found *KAT7* prefers histone substrates. p53 acetylation is only detectable in reactions without histones (Iizuka et al., 2008). To address whether the apoptosis associated with *KAT7* loss is mediated through p53 in AML, more p53 mutation cell lines need to be assayed. Possibly through generating isogenic p53 mutant cell lines of MOLM-13 or inhibiting *KAT7* in other p53 mutated AML cell lines.

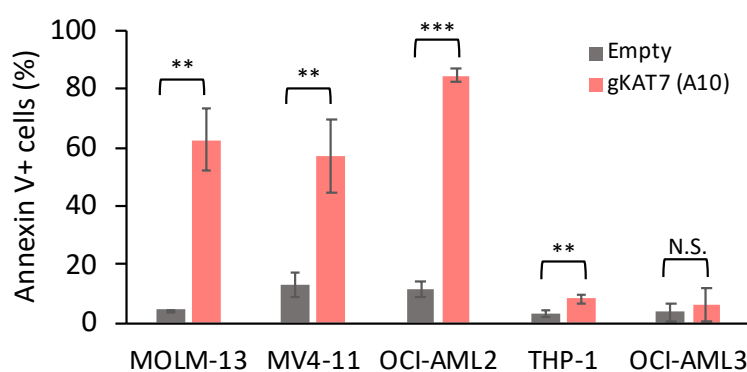


Figure 3.12 Apoptosis of *KAT7* knock-out in AML cell lines

AML cells were introduced with gKAT7(A10) via lentivirus transduction and BFP-positive cells were sorted day 3 post-transduction and stained with Annexin-V-APC antibody on day 9 post-transduction. Mean \pm S.D., $n = 3$ (biological replicates), two-tailed t -test (N.S. non-significant, $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$).

3.6.2 Proliferation

Next, we repeated the proliferation assay in the MOLM-13, MV4-11, THP-1, OCI-AML2, OCI-AML3 with gKAT7 (A5) in addition to gKAT7 (A10) that was previously used in Figure 3.5. Furthermore, we utilized HL-60 and Nomo-1 to investigate the essentiality of *KAT7* in other AML cell lines. The proliferation of HL-60 was not affected by the loss of *KAT7* whereas Nomo-1 proliferation was compromised in *KAT7* knock-out (Figure 3.13). Collectively, our findings so far suggest *KAT7* may be essential for MLL-fusion positive AML, but not in MLL-fusion negative cell lines.

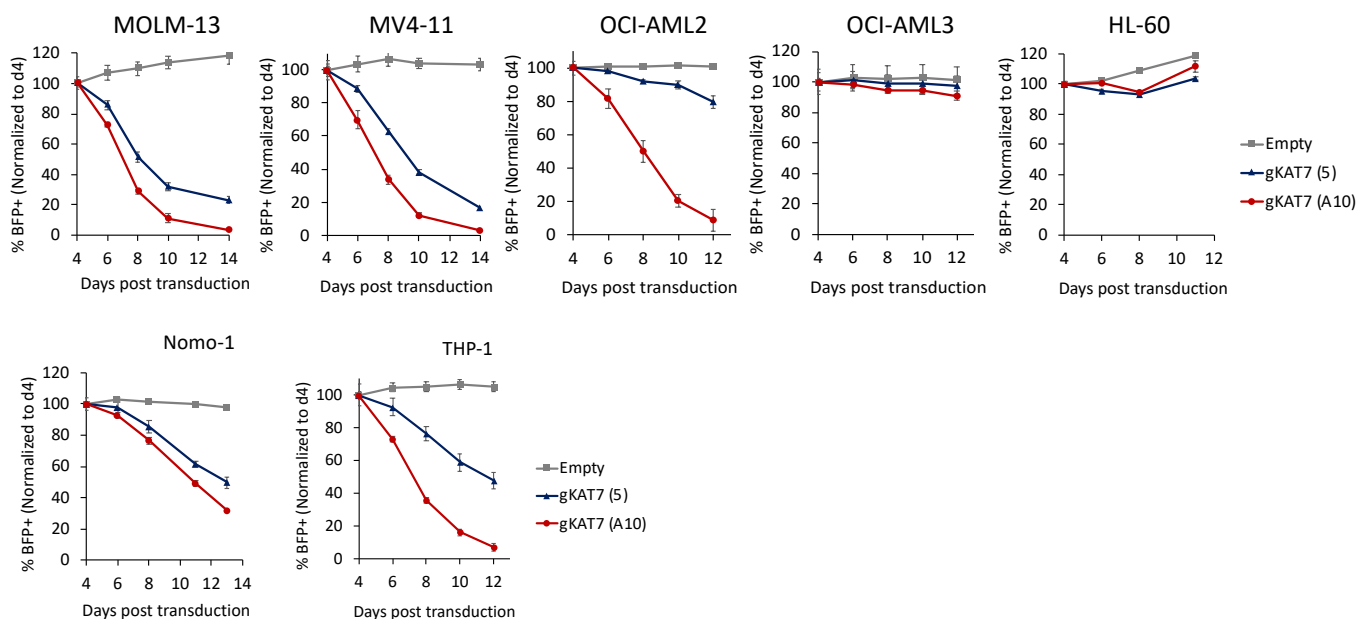


Figure 3.13 Proliferation of *KAT7* knock-out in AML cell lines

The proliferation of *KAT7* knock-out using two individual gKAT7 (5 and A10) or empty control in different AML cell lines. The relative percentage of BFP-positive (gKAT7-transduced) cells were assayed over time and all time points were normalized to day 4. Mean \pm SD, $n = 3$

In OCI-AML2, gKAT7 (5) transduced cells have a less pronounced effect on proliferation, compared to gKAT7 (A10) transduced cells of the same cell line. This may be due to the high transduction efficiency which resulted in ~80% transduced cells. Figure 3.14 illustrates a gRNA depletion simulation comparing the changes in proliferation of experiment starting with high transduction efficiency (80% cells transduced) with that of low transduction efficiency (30% cells transduced).

Assuming that non-transduced, BFP-negative cells double every 24 hours, whilst transduced, BFP-positive, KAT7 knock-out cells conferring growth disadvantage duplicate 1.8 times every 24 hours. An initial 80% transduced cells (20% non-transduced) at time point 0 will give rise to 43.9% viable cells at day 19. On the contrary, starting with 30% transduced cells (70% non-transduced) will result in 18.2% proliferative cells at day 19. Although KAT7 is an essential gene in some AML cell lines, such as OCI-AML2, the proliferation assay is based on the competition between BFP-positive and BFP-negative cells, therefore, the relative proportion of transduced and non-transduced cells can influence the dynamic of depletion. Keep transduction efficiency between cell lines and between gRNAs as similar as possible when making comparisons is ideal. Transduction efficiency is dependent on both virus titer as well as cell lines. Virus titer can vary between batches of virus production and titration for each batch is important for achieving consistent proportion of transduced cells at the start of every experiments.

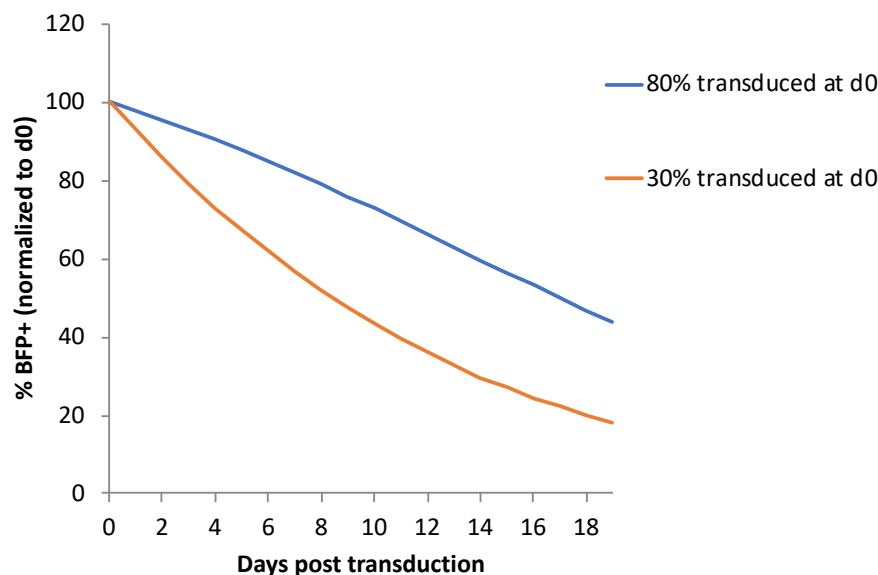


Figure 3.14 Simulation of cell proliferation with different percentage of initial transduced cells

Prediction of the cell proliferation dynamics of knock-out cells, conferring growth disadvantage, overtime with a high or low proportion of transduced cells on day 0. Given that the effect of knock-out is constant, high percentage of transduced cells (80% transduced at day 0) results in a slower depletion dynamic compared to a low percentage of transduced cells (30% transduced at day 0).

3.6.3 Differentiation

Differentiation time-course experiment in Chapter 3.5.2 showed day 7 represents the time point in which significant levels of CD11b marker was observed (Figure 3.8), we therefore assayed the effect of *KAT7* knock-out on CD11b expression using gKAT7 (A10) in various cell lines at this time point (Figure 3.15). Notably, an increase in CD11b was associated with the loss of *KAT7* in MOLM-13 (61%), OCI-AML2 (32%) and Nomo-1 (21.1%).

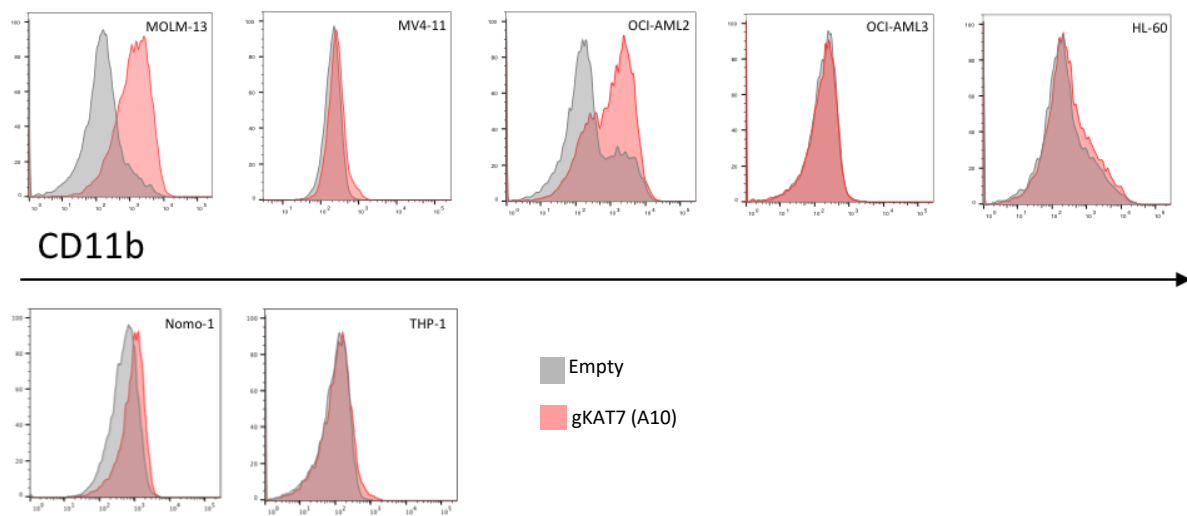


Figure 3.15 CD11b expression of *KAT7* knock-out in AML cell lines

*CD11b-APC staining in empty (grey) and *KAT7* knock-out (red) cells on day 7 post-transduction in different AML cell lines. BFP-positive knock-out cells were sorted day 3 post-transduction.*

Initially, we observed 40% population of differentiated cells in MV4-11 on day 6 (Figure3.16A) and by day 7, the CD11b-positive population reduced to 22.8% (Figure3.16B), suggesting that MV4-11 undergo apoptosis soon after signs of differentiation. However, this finding was not reproducible, therefore MV4-11 is unlikely to have increased CD11b expression following knock-out of *KAT7*. It is likely that following the loss of *KAT7*, MV4-11 induces apoptosis without differentiation. As we did not observe CD11b changes in THP-1 knock-out nor did we observe a drastic increase in apoptosis, we assessed CD14, another marker for myeloid differentiation, and day 7 showed 24.1% of differentiated cells (Figure3.16C).

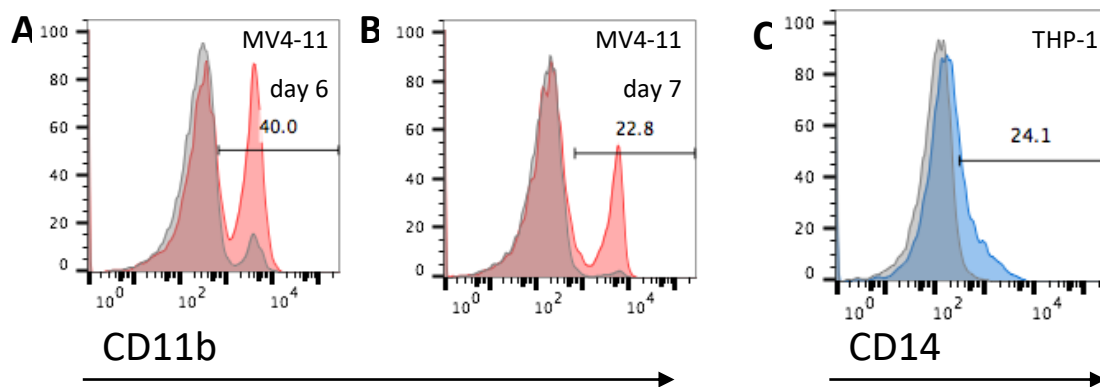


Figure 3.16 Expression of myeloid differentiation marker in KAT7 knock-out MV4-11 and THP-1

BFP-positive knock-out cells were sorted day 3 post-transduction. A and B) CD11b expressions in empty (grey) and KAT7 knock-out (red) MV4-11 cells on day 6 and day 7 post-transduction respectively. C) CD13 expressions in empty (grey) and KAT7 knock-out (blue) THP-1 cells on day 7 post-transduction.

3.7 In vivo effects of KAT7 loss

To assess if KAT7 inactivation has effects *in vivo*, we injected KAT7 knock-out luciferase-expressing MOLM-13 cells into NSGW41 male mice (derived by breeding NSG and KITW41 animal to homozygosity, i.e. $Kit^{W41/W41}$, $Prkdc^{-/-}$, and $Il2rg^{-/-}$ or $Il2rg^{+/Y}$). This breed of mice was chosen based on several considerations. Firstly, previous experiments have demonstrated efficient and sustained engraftment of MOLM-13 cells into the bone marrow of NSGW41 mice. Kit signalling is required for the expansion of progenitor cells, as such human Kit-proficient cells populate the bone marrow, replacing the endogenous murine Kit-mutant hematopoietic cells (Waskow et al., 2014). Secondly, the agouti coat of NSGW41 mice allows for better visualization of the bioluminescent signal, compared to dark-furred mice. Lastly, male mice were chosen as MOLM-13 was derived from a male AML patient. Mice transplanted with KAT7 knock-out MOLM-13 cells had significantly slower AML progression and increased survival compared to those injected with wild-type cells (Figure 3.17). Transplantation of KAT7 knock-out MOLM-13 cells eventually led to the development of abnormal masses at ectopic parts of the body, including under the skin and on the liver. This might be due to KAT7 knock-out cells undergoing differentiation, whereby the maturation of AML cells allows for “metastasis” in other organs. On the contrary, wild-type MOLM-13 localizes in the bone marrow of the mice.

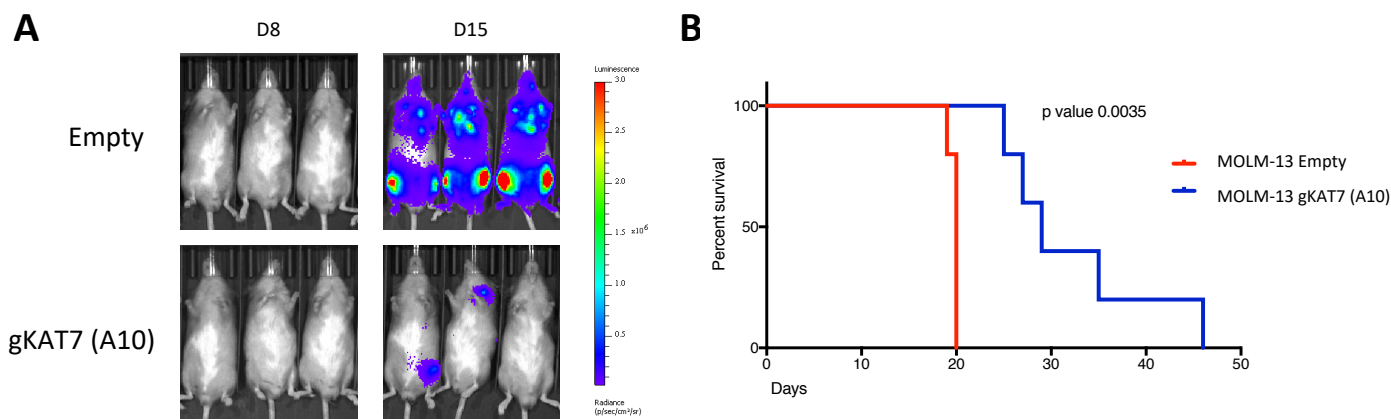


Figure 3.17 Disease progression and survival of mice injected with luciferase-expressing MOLM-13 cells

Luciferase expressing MOLM-13 with KAT7 knock-out or empty control cells were transplanted into NSGW41 male mice day 3 post-transduction. BFP-positive cells were sorted day 2 post-transduction. A) bioluminescent signal on day 8 and day 15 post-injection B) Kaplan-Meier survival plot (n=5 in each arm). Experiment and figures generated in collaboration with Etienne De Braekeleer.

3.8 Expression of KAT7 in normal haematopoiesis

To understand the expression of KAT7 in normal hematopoiesis, we utilized the BloodSpot platform (Bagger et al., 2019). Gene expression profiles of hematopoietic cells obtained from the bone marrow of healthy subjects showed that levels of KAT7 transcript fluctuate through hematopoietic differentiation (Bagger et al., 2019; Rapin et al., 2014). Notably, common myeloid progenitor (CMP) and megakaryocyte-erythroid progenitor (MEP) appear to have the highest expression of KAT7, compared to other hematopoietic cells in the myeloid lineage differentiation tree (Figure 3.18) (Bagger et al., 2019; Rapin et al., 2014), which may suggest that KAT7 is required to maintain the potency of these cells.

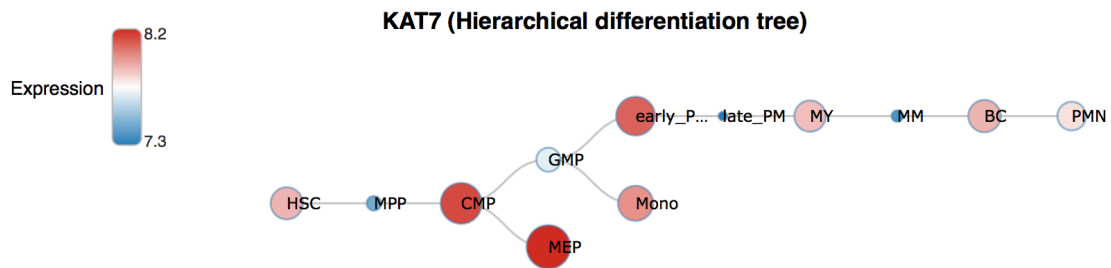
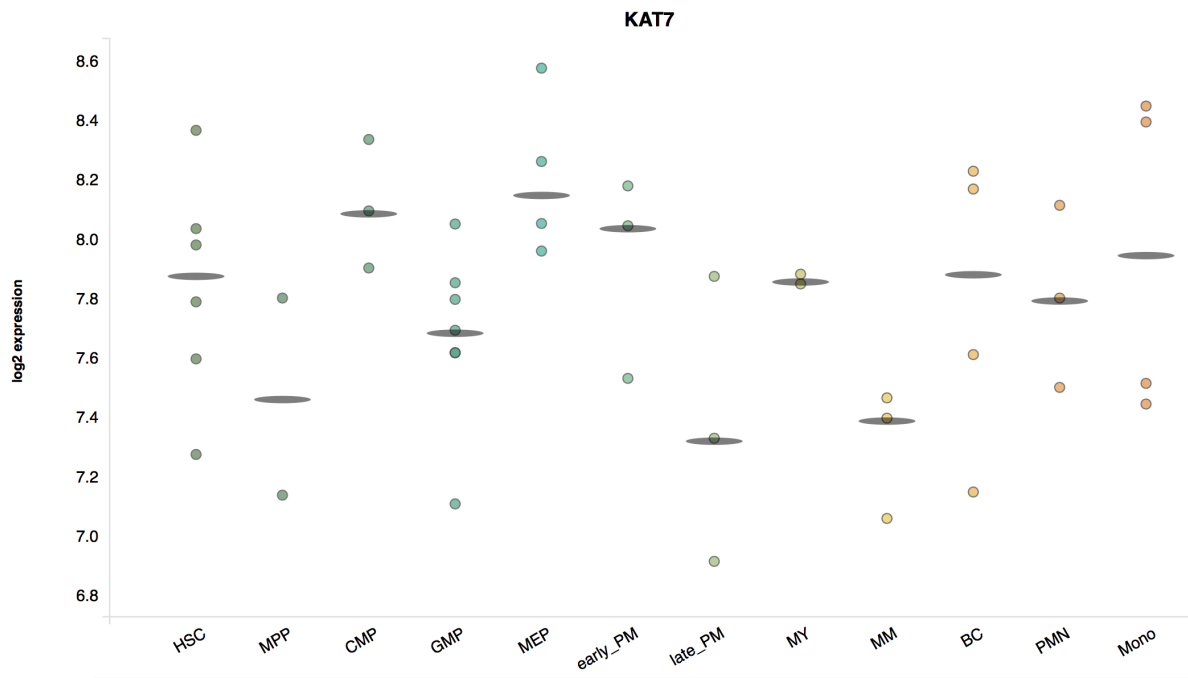
A**B**

Figure 3.18 Expression of KAT7 in normal hematopoietic cells

Transcripts (Log2 expression) of KAT7 in cells extracted from the bone marrow healthy patients (Bagger et al., 2019; Rapin et al., 2014). A) Myeloid hierarchical differentiation tree and expression of KAT7 in the corresponding stem and progenitor and myeloid cells. B) Jitter strip chart, corresponding to the cells depicted in the differentiation tree. Each dot represents one sample and the lines represent the median expression of each cell type. Hematopoietic stem cell (HSC); Multipotent progenitors (MPP); common myeloid progenitors (CMP); Granulocyte monocyte progenitors (GMP); megakaryocyte-erythroid progenitors (MEP); early promyelocyte (early_PM); late promyelocyte (late_PM); myelocyte (MY); metamyelocytes (MM); band cell (BC); polymorphonuclear cells (PMN); monocytes (mono).

3.9 KAT7 mutations and expression in other cancer types

Among normal cell types, KAT7 is most highly expressed in testis (Iizuka and Stillman, 1999; Sharma et al., 2000). In the context of cancer, analysis of the The Cancer Genome Atlas (TCGA) dataset (genomics data of 37 cancer types) showed that testicular germ cell cancer (TGCT) represents the tumour type with the highest expression of KAT7 (Figure 3.19) (Deng et al., 2017), possibly correlated to the high expression in normal testis cells. We utilized the cBioPortal platform that facilitates the analysis and visualization of large-scale cancer genomics datasets to look for mutations in KAT7 among AML patients. We assessed 3 AML genomic studies (TCGA Research Network, 2013; Tyner et al., 2018; Welch et al., 2016), which consist a total of 1008 AML patient samples (878 patients queried) and found no KAT7 alterations—mutations, amplification, deletion or fusion—among those profiled (Cerami et al., 2012; Gao et al., 2013a). The lack of KAT7 alterations found among AML patients could suggest that KAT7 is not an oncogene involved in AML initiation however, may play a critical role in maintaining leukemogenesis which could potentially be exploited therapeutically. Interestingly, we observed KAT7 mutations in other cancer types, most frequently in uterine corpus endometrial carcinoma (UCEC), whereby 7.89% of UCEC patients possess a mutation in this gene (Cerami et al., 2012; Gao et al., 2013a; Grossman et al., 2016). 164 mutations were detected among all the TCGA datasets, 118 of which are mutations in the protein-coding region of KAT7 (Grossman et al., 2016). Missense mutations are the predominant alterations in KAT7 among the various cancer types, constituting 89.8% (106 of 118) of all KAT7 alterations documented (Figure 3.20) (Grossman et al., 2016). There is no obvious mutational “hotspot” within the coding region of KAT7 and mutation sites appears to be distributed across the gene (Figure 3.20). It is unclear if these mutations are oncogenic and contribute to the fitness of other cancer types.

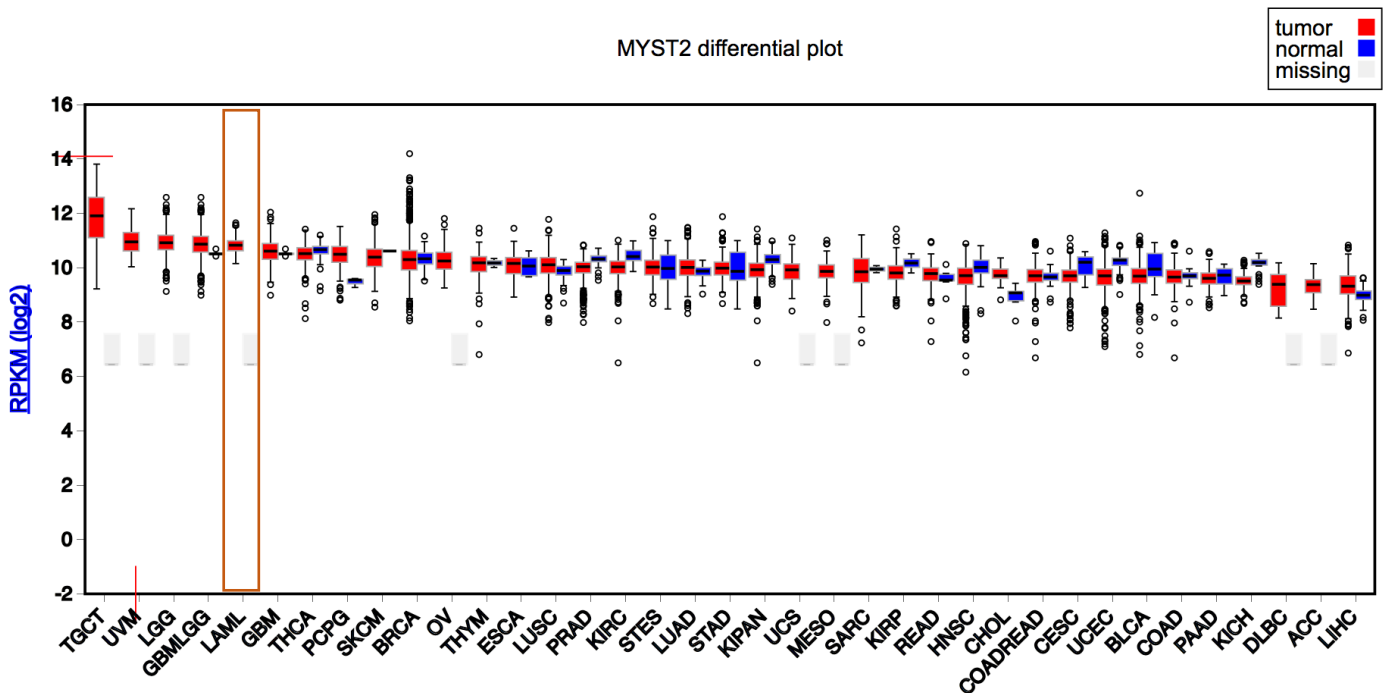


Figure 3.19 KAT7 expression in 37 cancer types from The Cancer Genome Atlas (TCGA)

Log2 transformed reads per kilobase million (RPKM) transcript levels of KAT7 obtained from FireBrowse (Deng et al., 2017). In the corresponding tumour types (red), as well as normal tissues (blue). Normal tissue samples missing in some cancer types (grey). Testicular germ cell tumour (TGCT); Uveal melanoma (UVM); brain lower grade glioma (LGG); glioma (GBMLGG); acute myeloid leukemia (LAML); glioblastoma multiforme (GBM); thyroid carcinoma (THCA); Pheochromocytoma and Paraganglioma (PCPG); skin cutaneous melanoma (SKCM); breast invasive carcinoma (BRCA); ovarian serous cystadenocarcinoma (OV); thymoma (THYM); esophageal carcinoma (ESCA); lung squamous cell carcinoma (LUSC); prostate adenocarcinoma (PRAD); kidney renal clear cell carcinoma (KIRC); stomach and esophageal carcinoma (STES); lung adenocarcinoma (LUAD); stomach adenocarcinoma (STAD); Pan-kidney cohort (KIPAN); uterine carcinosarcoma (UCS); mesothelioma (MESO); sarcoma (SARC); kidney renal papillary cell carcinoma (KIRP); rectum adenocarcinoma (READ); head and neck squamous cell carcinoma (HNSC); Cholangiocarcinoma (CHOL); colorectal adenocarcinoma (COADREAD); cervical and endocervical cancer (CESC); bladder urothelial carcinoma (BLCA); colon adenocarcinoma (COAD); pancreatic adenocarcinoma (PAAD); kidney chromophobe (KICH); lymphoid neoplasm diffuse large b-cell lymphoma (DLBC); adrenocortical carcinoma (ACC); liver hepatocellular carcinoma (LIHC).

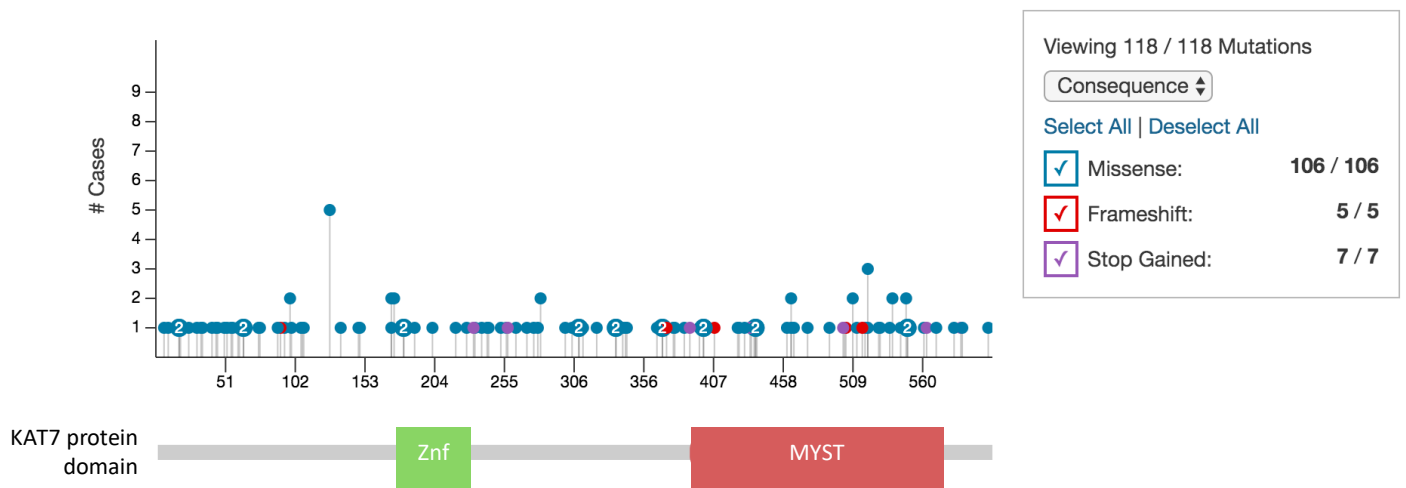


Figure 3.20 Location and type of mutations in the protein-coding region of KAT7 found across The Cancer Genome Atlas (TCGA) dataset

Image obtained from NCI GDC Data Portal (Grossman et al., 2016). Each circle represents a mutation. The position of each mutation is presented relative to the domains of KAT7 protein. The height (y-axis) of the circles indicates the frequency of the mutation within the TCGA dataset. Missense mutations (blue), frameshift mutations (red), stop gained mutations (purple) are indicated by the respective colours. The number within a circle represents the number of different mutations found at the same position.

3.10 Discussion

CRISPR-Cas9 dropout screens provide a wealth of novel therapeutic targets. Validation of genome-wide CRISPR-screens is imperative due to false-positive and false-negative hits. Despite the 8 chosen novel druggable candidate genes are significant dropouts (<20% FDR) in the MOLM-13 cell line, only *SIK3*, *KAT6A* and *KAT7* knock-outs show strong phenotypic consequence on cellular proliferation over the course of 14 days. Whilst *KDM5B* knock-out shows no effect in MOLM-13, *HIPK1*, *STK3*, *RSK1* and *KDM2A* knock-outs induce a weak effect on the proliferative ability of MOLM-13 cells based on our validation. This may be partly explained by the difference in the duration of the experiment between validation (14 days) and screening (30 days). These genes may affect the fitness of MOLM-13 over a longer time-course. Although it is possible that the chosen gRNAs for each gene may not effectively modify the genetic target, thus result in functional protein, this scenario is less likely given our gRNA selection criteria were based on those that were depleted in the AML CRISPR-Cas9 screen performed by Tzelepis et al. From our validation, we found *SIK3* is additionally essential

in MV4-11 and OCI-AML2 as this was not picked up from the screens. The output of *KAT7* validation is consistent with the screening output across all AML cell lines. Interestingly, *KAT5* is also a member of the MYST family protein like *KAT6A* and *KAT7* and is an AML-specific essential gene (Tzelepis et al. screening results reflect dropout in MOLM-13 and OCI-AML2 cell lines), but was not identified as a druggable gene in the DGIdb characterization.

SIK3 knock-out has a strong defect in proliferation across multiple AML cell lines. However, this phenotype was not associated with a significant increase in apoptosis or differentiation. The proliferation may be associated with growth or cell cycle arrest. Recent studies have shown that the catalytic activity of *SIK3* is important for the proliferation of MLL-fusion positive AML cells (Tarumoto et al., 2018). Tarumoto et al., went on to show that *SIK3*-dependent phosphorylation of HDAC4, prevents HDAC4-mediated inhibition of MLL-fusion target genes such as *MEF2C*.

Genetic ablation of *KAT7* induces anti-proliferation and pro-apoptotic phenotypes in various AML cell lines, particularly those harbouring MLL-fusion mutations. Myeloid differentiation of various magnitude is further observed in *KAT7* knock-out in the MLL-fusion positive AML cell lines MOLM-13 (MLL-AF9), OCI-AML2 (MLL-AF6), THP-1 (MLL-AF9) and Nomo-1 (MLL-AF9) but not in MV4-11 (MLL-AF4). Before the knock-out cells eventually undergo cell death, cell lines that differentiate after *KAT7* knock-out may temporarily restore the function of the transformed hematopoietic cell. Our validation revealed *KAT7* is a genetic vulnerability in an MLL-AF4 positive AML cell line. MLL-AF4 translocation is commonly associated with acute B lymphoblastic leukaemia (B-ALL) (Ayton and Cleary, 2001; Meyer et al., 2013, 2018). *KAT7* may also be essential for MLL-AF4 positive B-ALL and whether *KAT7* can be therapeutically exploited in this subtype of leukemia is an interesting avenue to explore.

We have briefly investigated the effects of *KAT7* loss *in vivo*, by transplanting *KAT7*-knock-out luciferase-expressing MOLM-13 cells into immunodeficient mice. Even though this experiment showed that mice harbouring MOLM-13 *KAT7* knock-out have delayed disease progression and prolonged survival, the *in vitro* studies thus far suggests that differentiation may be initiated by the day of injection (day 3 post-transduction) and could affect the engraftment of AML cells at the bone marrow. Cells undergo differentiation soon after transduction (approximately 30% CD11b differentiated cells on day 5 post-transduction), and thus would likely to have changes in gene expression by day 3 post-transduction, thereby changing the properties of the cells. This could explain the ectopic engraftment of *KAT7* knock-out MOLM-13 cells seen in the mice experiment.

An inducible *KAT7*-knock-out mouse model would be invaluable to comprehensively assess the loss of *KAT7 in vivo*, particularly in evaluating if the loss of *KAT7* has detrimental effects. Furthermore, the lack of *KAT7* inhibitor presently means that it is not yet possible to assess the wholistic effect of inhibiting *KAT7* at an organismal level, including effects on normal haematopoiesis and other cell types. Development of small molecules or drugs against *KAT7* will facilitate our understanding of whether targeting *KAT7* results in minimal toxicity and possess AML-specific efficacy. Encouragingly, a recent genome-wide CRISPR-Cas9 study on haematopoiesis of human primary CD34+ human progenitor and stem cells (HPSCs) and T-cells (Ting et al., 2018) did not identify *KAT7* as a significant player in the development of these cells, which suggests that inhibiting *KAT7* may not have detrimental effects on the development of normal blood cells. Possible additional experiments using the CRISPR-Cas9 based gene-editing methods include knocking-out *KAT7*, by co-transfection of gRNA and Cas9, in CD34+ human cord blood cells and AML cells (MLL-fusion positive and MLL-fusion negative) from human patients.

Chapter 4: Molecular characterization of KAT7 in AML

4.1 Introduction

In the previous chapter, we validated KAT7 as an essential gene in AML following reports of it being a CRISPR-Cas9 dropout candidate (Tzelepis et al., 2016; Wang et al., 2017). Histone acetyltransferases can be found in many eukaryotic organisms, ranging from yeast to human. This class of enzyme controls cellular homeostasis and a plethora of biological processes through their function in catalysing the transfer of acetyl group from acetyl-CoA to ϵ -lysine residues onto proteins, including histones proteins (Sheikh and Akhtar, 2018). Among many important functional roles of KAT7 is its ability to regulate gene expression (Avvakumov et al., 2012; Saksouk et al., 2009; Sheikh and Akhtar, 2018; Yan et al., 2018).

Information on KAT7 in the context of AML is scarce with only one study to date that has explored this topic. Sauer and colleagues reported that the expression of KAT7 is decreased in human AML blast cells compared to the non-malignant CD34+ progenitor cells. Furthermore, the overexpression of *KAT7* negatively affects the proliferative ability of mouse hematopoietic stem cells and the knock-down of *KAT7* by shRNA boosted colony formation in THP-1 (AML cell line) and SEM (B-ALL cell line) (Sauer et al., 2015). The inconsistency between this study and our findings thus far prompted us to explore further into the molecular function of KAT7 in our cellular models of human AML.

In this chapter, we characterize the changes in histone acetylation marks following KAT7 loss and explore whether the catalytic function of KAT7 is important for sustaining AML. We investigate the genome-wide binding profiles of KAT7 to identify its distribution across various functional elements of protein-encoding genes. We subsequently link KAT7 occupancy with gene expression data to narrow down the group of genes that are likely regulated by KAT7. Together, the aim is to enhance our understanding of the general molecular functions of KAT7 in AML cells.

4.2 KAT7 protein and H3K14ac levels are depleted after gKAT7 mediated knock-out

In chapter 3, we concluded that introducing KAT7 gRNA resulted in proliferation, apoptosis and/or differentiation phenotypes in many AML cell lines. To determine if these phenotypes are attributed to on-target editing in the KAT7 gene by the selected gKAT7, we assess if KAT7 is knocked-out at a

protein level using gKAT7 (A10). Firstly, we wanted to establish the time point at which KAT7 protein is depleted after transduction. We assessed day 3 and day 4 post-transduction in MOLM-13 and THP-1 and found that KAT7 protein in gRNA transduced cells are reduced, compared to empty control, but still detectable at day 3 for both cell lines (Figure 4.1A). By day 4, KAT7 protein levels are abolished for both cell lines (Figure 4.1A). We further assess H3K14ac, a histone mark modified by KAT7 and strongly reduced with the loss of KAT7 in vivo (Kueh et al., 2011). Interestingly, global H3K14ac levels are depleted in knock-out cells compared to empty control at day 3 post-transduction, even in the presence of detectable levels of KAT7 protein (Figure 4.1A). Expectedly, levels of this histone mark are further reduced at day 4 (Figure 4.1A). These data, therefore, reveal that at day 4 for post-transduction, the global KAT7 protein level is abolished and the loss of KAT7-modified histone marks is rapid. Importantly, we detected the on-target effect by gKAT7 (A10) and day 4 is an appropriate time point for further analysis.

4.3 KAT7 modified histone marks H3K14ac and H4K12ac are depleted in *KAT7* knock-out cells

In vitro and in vivo studies have elucidated the histone lysine residues acetylated by KAT7, as mentioned in Chapter 1. Briefly, KAT7 protein complex has specificity for lysine residues on histone 3 and histone 4 tails. On histone 3, lysine residues 14 and 23 are reported targets for KAT7-BRPF complex (Feng et al., 2016; Kueh et al., 2011; Lalonde et al., 2013; Mishima et al., 2011). On histone 4, KAT7-JADE complex acetylates lysine residues 5, 8 and 12 (Doyon et al., 2006; Iizuka et al., 2009). We hence investigate whether loss of KAT7 results in changes in the levels of these histone lysine acetylation, using MOLM-13 as the model cell line. Strikingly, we see a global loss in H3K14ac and H4K12ac acetylation marks in the knock-out, relative to the control, on day 4 after transduction (Figure 4.1B). Minor depletion of H4K8ac marks can be seen in the knock-out relative to the empty control (Figure 4.1B) and H3K23ac and H4K5ac appear to have no global change associated with the loss of KAT7 protein (Figure 4.1B). These data suggest that KAT7 is the main HAT responsible for global H3K14ac and H4K12ac levels in MOLM-13, whilst the levels of other known KAT7 histone targets may be compensated or acetylated by other HATs within the cell.

Next, we investigated if levels of H3K14ac and H4K12ac are reduced in other AML cells, including cell lines which proliferation was unaffected by *KAT7* knock-out. Indeed, across various AML cell lines, global depletion of H3K14ac and H4K12ac was observed, irrespective of whether *KAT7* is an essential gene in that cell line (Figure 4.1C). Although we cannot attribute the loss of H3K14ac and H4K12ac

marks to the proliferation and apoptosis phenotype from these set of experiments, however, we can deduce that H3K14ac and H4K12ac are dispensable for general proliferation and apoptosis. This is because, despite the loss of KAT7 protein and the corresponding depletion of these two histone marks, OCI-AML3 and HL-60 can proliferate and not induce apoptosis.

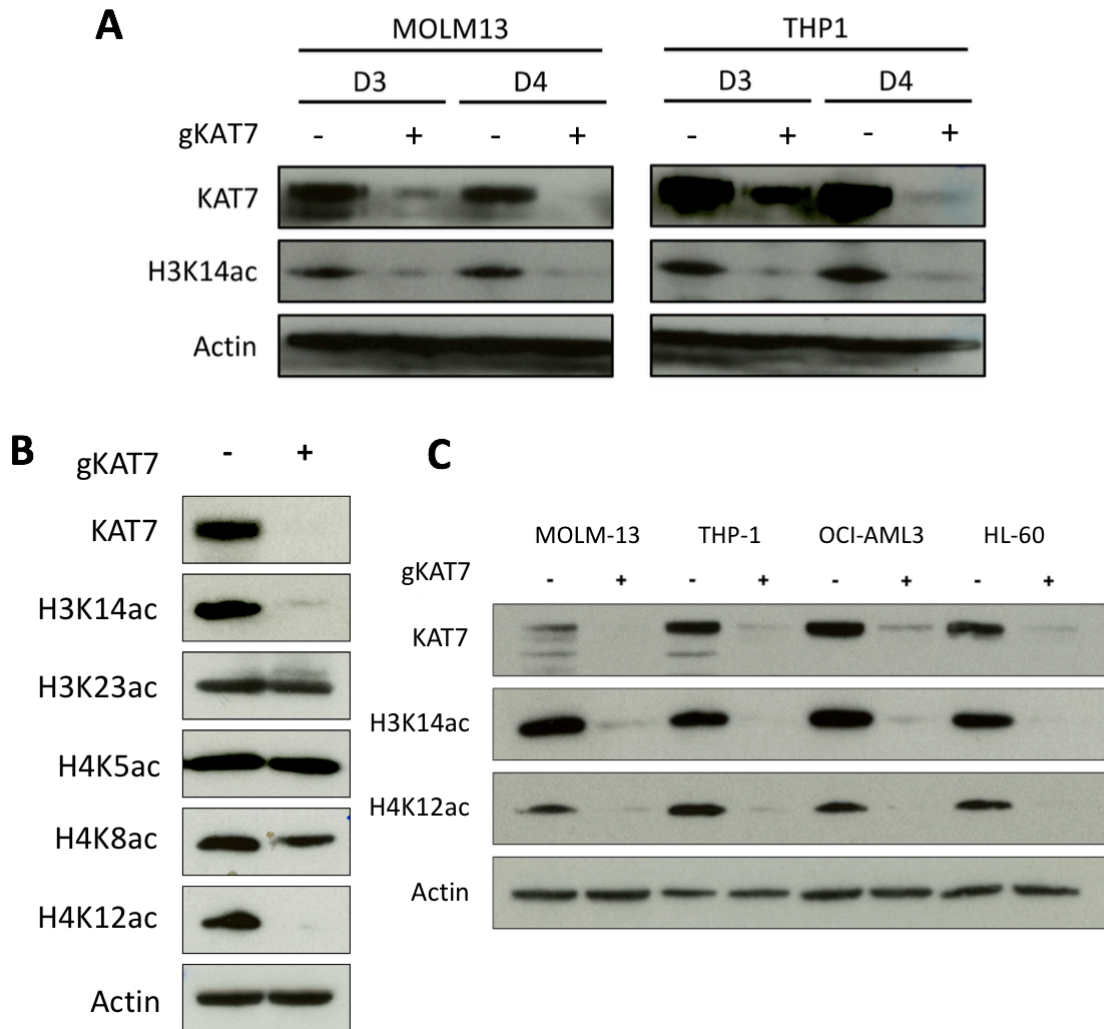


Figure 4.1 Changes in levels of KAT7 protein and KAT7-mediated histone lysine acetylation following gRNA-mediated knock-out of KAT7

AML cell lines were transduced with lentivirus expressing empty (-) or gKAT7(A10) (+) and were sorted by flow cytometry for BFP-positive population on day 3 post-transduction. A) Time-course experiments in MOLM-13 and THP-1. Lysates were harvested on day 3 (D3) and day 4 (D4) post-transduction for immunoblot of KAT7 and H3K14ac. B) Immunoblots of all known KAT7 histone lysine acetylation targets on day 4 post-transduction in MOLM-13. C) Changes in global H3K14ac and H4K12ac levels in MLL-fusion positive (MOLM-13 and THP-1) and MLL-fusion negative (OCI-AML3 and HL-60) cell lines on day 4 post-transduction.

4.4 Changes in other histone modifications upon KAT7 loss

To begin to address if the global level of gene expression is affected after *KAT7* knock-out, we looked into various histone modifications in MOLM-13 on day 4 post-transduction. We first looked at histone modifications that are found on the promoter region actively transcribed genes such as trimethylation at H3K4 (H3K4me3) (Barth and Imhof, 2010; Liang et al., 2004; Zhang et al., 2015) and found that knock-out of *KAT7* did not lead to changes in global H3K4me3 (Figure 4.2A). We next assessed H3K9ac, which is commonly found to be positively correlated with both gene expression and H3K14ac levels (Barth and Imhof, 2010; Karmodiya et al., 2012; Liang et al., 2004). As histone 3 lysine 9 residue is not a known direct target for *KAT7*, it is therefore not surprising to observe no changes in the acetylation level with the loss of *KAT7* protein (Figure 4.2A).

H3K27ac, a histone modification that is also associated with the promoter region of active genes as well as with enhancers (Creyghton et al., 2010; Heintzman et al., 2009; Sheikh and Akhtar, 2018; Wang et al., 2008), decreases after *KAT7* loss (Figure 4.2A). This is an interesting observation as it implies that the loss of H3K14ac and/or H4K12ac may result in reduced acetylation levels of neighbouring lysine residues. It could be that *KAT7*-mediated acetylation can act as docking sites for other HATs, such as p300 and CBP that writes H3K27ac (Bose et al., 2017; Jin et al., 2011; Sheikh and Akhtar, 2018; Tie et al., 2009). To probe into this further, we performed a time course experiment and observed that like H3K14ac, H3K27ac is gradually reduced but not abolished at day 5 post-transduction in MOLM-13 cell line (Figure 4.2C). Furthermore, from day 3 after transduction in which we see the depletion of global H3K14ac levels, we do not observe as much decrease in H3K27ac, when comparing the knock-out with the empty control. This further implies that the loss of H3K27ac is a secondary event following loss of *KAT7* protein and *KAT7* modified acetylation marks. This change in global H3K27ac levels is not as pronounced in other AML cells such as THP-1 and OCI-AML3 on day 4 post-transduction (Figure 4.2B), compared to MOLM-13 (Figure 4.2C), suggesting that the magnitude of change in H3K27ac levels may be cell-line specific.

In addition, we also assessed levels of H3K27me3, a transcriptional repressive mark at promoters (Barth and Imhof, 2010; Calo and Wysocka, 2013; King et al., 2016; Wang et al., 2008) and saw no apparent change of this histone modification globally in MOLM-13 *KAT7* knock-out cells on day 4 post-transduction (Figure 4.2A). Similarly, the levels of H3K79me2, a mark that is only known to be methylated by DOT1L and is associated with the gene body of actively-transcribed genes (Barth and Imhof, 2010; Nguyen and Zhang, 2011; Steger et al., 2008; Wood et al., 2018), did not change in knock-out cells compared to empty control (Figure 4.2A). The turnover rates and half-life of

methylation marks are relatively slower and longer respectively, compared to histone acetylation and phosphorylation (Barth and Imhof, 2010), therefore assessing at a later time point might be more appropriate. Conversely, later time point changes are more likely to be secondary effects as opposed to acute, primary consequences. Collectively, these findings suggest that loss of acetylation marks, particularly H3K14ac and H4K12ac are the primary changes following the loss of KAT7 protein in AML cells.

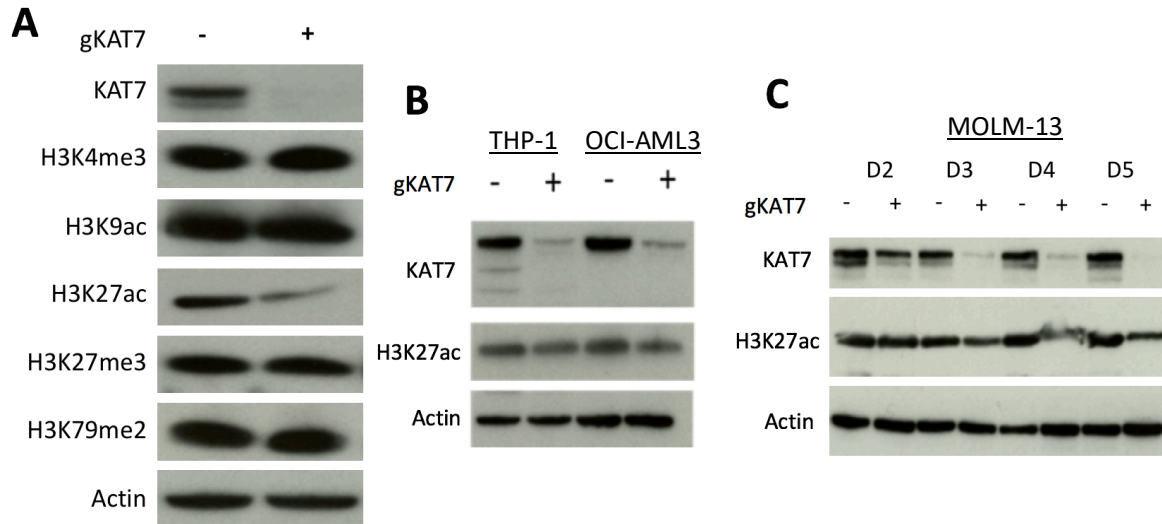


Figure 4.2 Levels of activating and inactivating histone marks after KAT7 knock-out

AML cell lines were transduced with lentivirus expressing empty (-) or gKAT7(A10) (+) and cells were sorted by flow cytometry for BFP-positive population on day 2 or day 3 post-transduction. A) Immunoblot of activating histone marks H3K9ac, H3K27ac, H3K4me3 and H3K79me2; and inactivating histone marks H3K27me3 and H3K27ac of MOLM-13 lysates harvested on day 4 post-transduction. B) Immunoblot of H3K27ac levels in THP-1 and OCI-AML3 on day 4 post-transduction. C) Time course experiment in MOLM-13. Lysates were harvested from day 2 to day 5 post-transduction for the immunoblot of H3K27ac.

4.5 KAT7 catalytic activity

We next investigated whether H3K14ac and H4K12ac are associated with the previously observed proliferation and differentiation phenotypes in cell lines in which KAT7 is an essential gene. To

address this, we can ask if the catalytic activity of KAT7 is required for the phenotypes. Interestingly, among the gKAT7 used in the study, gRNAs that target the MYST catalytic domain –gKAT7 (A10) and (5)—resulted in the most drastic proliferative defect, compared to gRNAs that target the serine-rich domain (gKAT7 (1) and (2)), zinc finger domain (gKAT7 (3)) or gRNA that does not target a protein domain (gKAT7 (4)) (Figure 4.3 and Figure 3.6A proliferation in chapter 3). gRNAs targeting of the non-catalytic domain may result in truncated but catalytic active protein product, whereas gRNAs that target the highly conserved MYST domain are likely to impair the catalytic function even if truncated KAT7 proteins are produced.

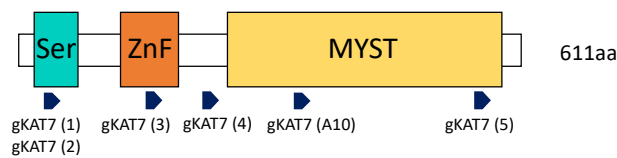


Figure 4.3 Schematic of KAT7 protein and the targeting sites of the different gKAT7

The various gKAT7 used in this study, targets different protein domains of KAT7. gKAT7 (1) and gKAT7 (2) targets DNA sequences corresponding to the Serine-rich domain (Ser) while gKAT7 (3) targets the Zinc finger domain (ZnF) of KAT7. Sequences modified by gKAT7 (4) do not correspond to a protein domain. gKAT7 (A10) and gKAT7 (5) target the MYST catalytic domain and display the most drastic proliferative defect.

4.5.1 G485A catalytic active mutant

We initially adopted the G485A MYST domain mutation of KAT7 that is widely referenced in the literature (Chen et al., 2013; Iizuka et al., 2008, 2009, 2013; Miotto and Struhl, 2010; Wong et al., 2010) to investigate the involvement of KAT7 catalytic activity in the observed phenotype, namely proliferation and differentiation, in MOLM-13. As *KAT7* is an essential gene, we utilized a knock-in-knock-out approach to first introduce the G485A mutant *KAT7* transgene followed by knock-out of the endogenous *KAT7*. Importantly, the *KAT7* transgene is resistant to gRNA-mediated modifications by gKAT7 (5 and A10), by altering the nucleotide sequences to create synonymous mutations in the gRNA targeted region of the *KAT7* cDNA. Therefore, the endogenous *KAT7* gene can be ablated without affecting the mutant *KAT7* transgene.

The KAT7 construct is designed to incorporate a selectable marker for the positive integration and expression of mutant KAT7. To achieve this, we link the GFP transgene to the KAT7 transgene using T2A sequence, where both genes are under the expression of EF1- α promoter (Figure 4.4A). GFP and KAT7 transgenes are transcribed and translated as one mRNA transcript. Once translated, the T2A peptide induces “self-cleavage”, generating two separate proteins (Ibrahimi et al., 2009; Kim et al., 2011)—GFP and mutant KAT7. The levels of GFP signal, therefore, reflects the integration of KAT7 transgene into the genome and the expression of KAT7 mutant protein. The targeted G485A mutation was confirmed at the plasmid level by Sanger Sequencing (Figure 4.4B). Using the same approach described above, we also generated wild-type KAT7 that is resistant to gKAT7 (5 and A10). The KAT7 transgene was subsequently introduced to MOLM-13 cells via lentivirus transduction. The resulting MOLM-13 cells express the transgene KAT7 (G485A or wild-type) and the endogenous KAT7, hereafter referred as G485A KAT7 and wild-type KAT7 MOLM-13.

To assess the function of the G485A KAT7 mutant, we targeted the endogenous *KAT7* by gRNA-mediated knock-out so that the MOLM-13 cells only express the mutant KAT7. To our surprise, we found that G485A KAT7 mutant did not result in loss of H3K14ac (Figure 4.5A). This finding is inconsistent in two aspects. Firstly, previous reports of G485A implied that this mutation in KAT7 results in abolished of catalytic activity (Chen et al., 2013; Iizuka et al., 2008, 2009, 2013; Miotto and Struhl, 2010; Wong et al., 2010). Secondly, based on our findings above and findings of previous in vivo studies which implied the loss of H3K14ac is associated with the loss of KAT7 protein (Kueh et al., 2011; Mishima et al., 2011), we would expect H3K14ac to be abolished or at least strongly reduced in a KAT7 catalytic dead mutant. A catalytic mutant of KAT7 would predictively phenocopy the effects of KAT7 protein loss if KAT7 enzymatic activity is required for the acetylation of histone 3 lysine 14. Proliferation and differentiation assays also showed that G485A KAT7 mutant does not affect the proliferative potential of MOLM-13 cells (Figure 4.5B) and does not induce CD11b expression (Figure 4.5C). In contrast, the GFP control MOLM-13 cell lines phenocopies parental MOLM-13 in both proliferation and differentiation phenotypes (Figure 4.5B-C).

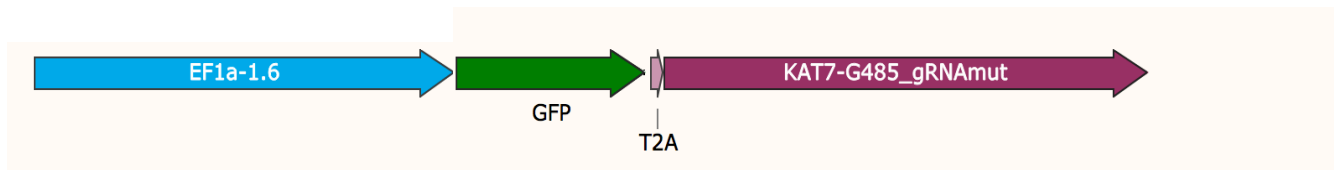
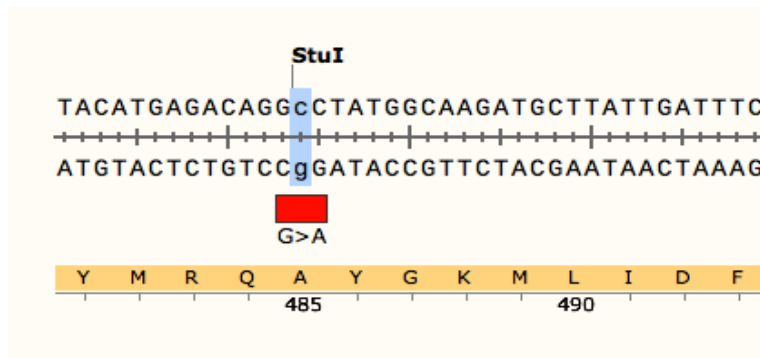
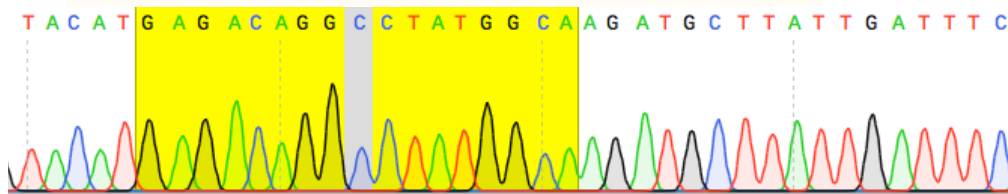
A**B****C**

Figure 4.4 Genetic construct of G485A mutant KAT7 transgene.

A) Expression of G485A KAT7 mutant under the expression of EF1a promoter. GFP is expressed under the same promoter and separated with the G485A transgene by a T2A peptide. KAT7 cDNA is resistant to gKAT7 (5) and gKAT7 (A10) modification by generating synonymous mutations in the DNA sequence at the gRNA targeting sites. B) Single nucleotide mutagenesis (highlighted in blue) to generate the glycine to alanine (G>A) substitution at amino acid position 485. C) Validation of G485A KAT7 by Sanger sequencing after cloning of KAT7 cDNA into the plasmid backbone.

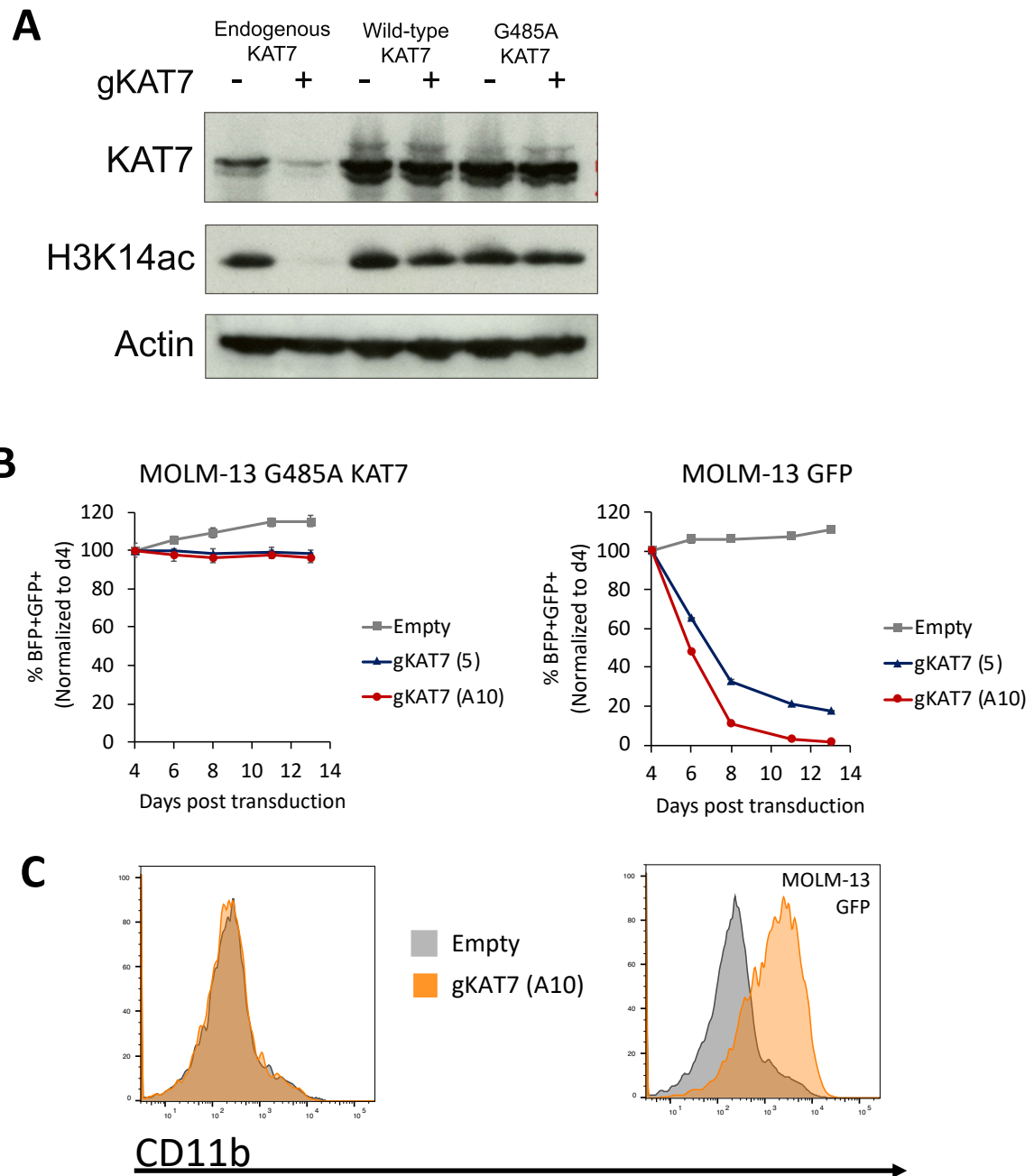


Figure 4.5 G485A mutant possesses catalytic activity and does not affect proliferation and differentiation of MOLM-13

A and C) MOLM-13 cell lines were transduced with lentivirus expressing empty (-) or gKAT7(A10) (+) and cells were sorted by flow cytometry for BFP-positive population on day 3 post-transduction. A) Parental MOLM-13 with GFP-expression (endogenous KAT7) and transgenic MOLM-13 cells with either knock-in of wild-type KAT7 (Wild-type KAT7) or G485A mutant (G485A KAT7). Changes in H3K14 acetylation levels day 4 post-transduction using gKAT7 (A10) to knock-out endogenous KAT7 but not wild-type or G485A mutant KAT7 transgene. B) The proliferation of MOLM-13 cells with

endogenous KAT7 knock-out in the presence of G485A mutant (MOLM-13 G485A KAT7) compared to parental MOLM-13 cell line (MOLM-13 GFP). The relative percentage of BFP and GFP double-positive cells were assayed over time and all the time points were normalized to day 4. Mean \pm S.D., n= 3. C) Left panel: CD11b staining of MOLM-13 cells expressing GFP, G485A KAT7 and no endogenous KAT7 (orange) and in cells expressing GFP, endogenous KAT7 and G485A KAT7 (grey); right panel: CD11b staining of parental MOLM-13 cells expressing GFP and endogenous KAT7 (grey) and of cells with KAT7 knock-out (orange), assayed on day 7 post-transduction.

4.5.2 E508Q catalytic dead mutant

We went on to search for other catalytic mutants by looking into other residues to mutate in the MYST domain of KAT7. Information on catalytic dead mutants in human KAT7 protein is scarce, however, MYST domain appears to be evolutionarily conserved between human and *Saccharomyces cerevisiae* (Figure 4.6) and the enzymatic activity of Esa1 in budding yeast has been extensively studied. Figure 4.6 shows the sequence homology of the human KAT7 protein with the mouse KAT7 and the budding yeast Esa1. The boxed area in Figure 4.6 emphasizes the region of the MYST domain where mutation(s) of the highlighted amino acid has been shown to disrupt the enzymatic activity (Adachi et al., 2002; Berndsen et al., 2007; Decker et al., 2008; Smith et al., 1998; Yan et al., 2002; Yuan et al., 2012). The amino acid highlighted in blue was the one we chose to mutate in our study as the glutamic acid (E) to glutamine (Q) substitution in Esa1 give rise to a charge-switch mutation, thereby inducing maximal disruption to the catalytic activity. This residue corresponds to the amino acid at position 508 (E508Q) of the human KAT7 protein and the nucleotide sequence that give rise to this mutation is depicted in Figure 4.7A. Figure 4.7A also illustrates the synonymous mutations (lower case nucleotides) that render the E508Q mutant KAT7 transgene unmodifiable by gKAT7(A10) and gKAT7(5). The selectable marker for the integration and expression of E508Q mutant transgene is identical to one described above for G485A mutant and also introduced to MOLM-13 cells via lentivirus transduction. Cloning of E508Q was confirmed by Sanger Sequencing at the plasmid level (Figure 4.7B).

[illegible]

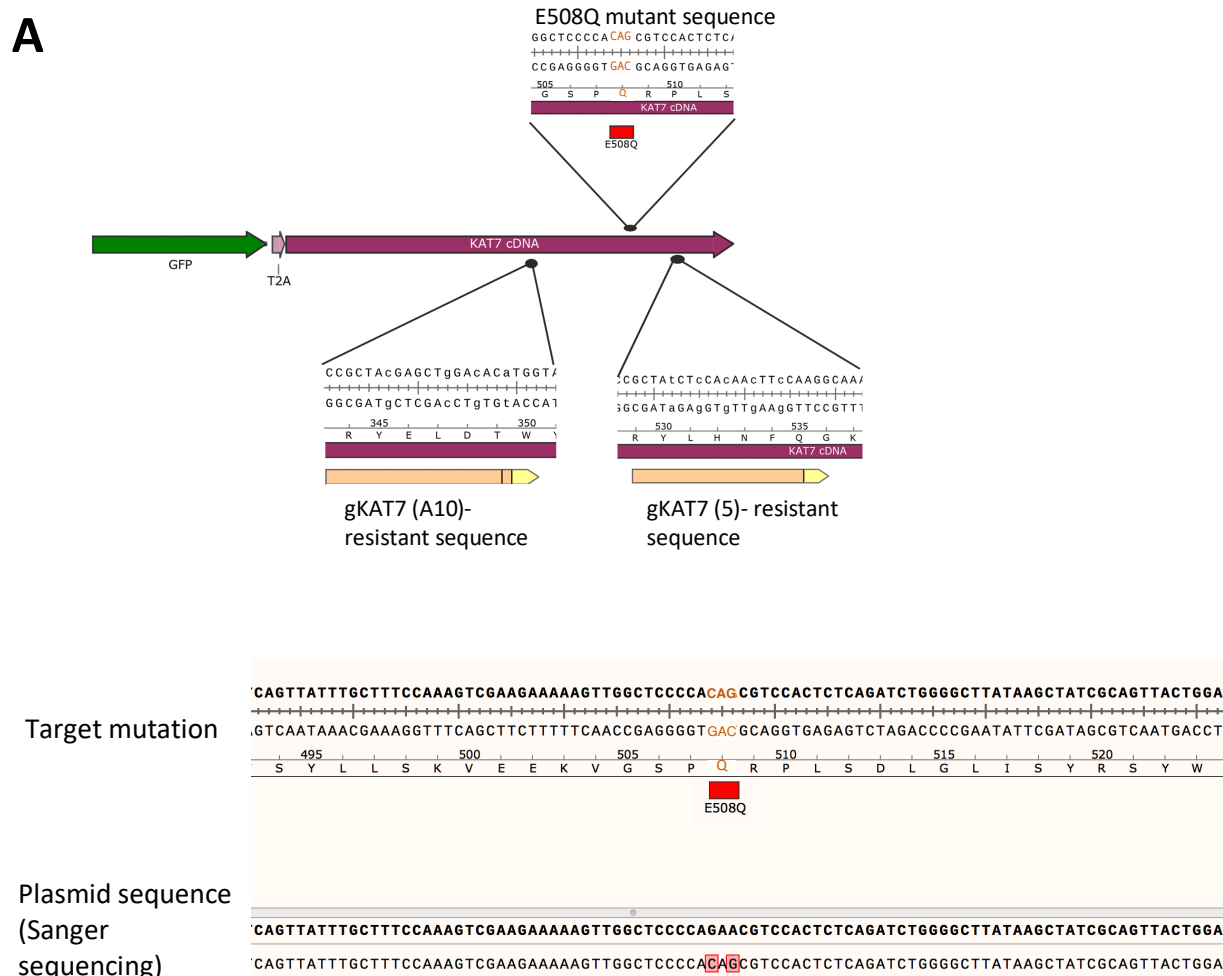


Figure 4.7 Genetic construct of E508Q mutant KAT7 transgene.

A) Knock-in transgene construct of KAT7 cDNA harbouring the glutamic acid to glutamine (E>Q) substitution and GFP separated by T2A peptide. Nucleotide sequences that give rise to the E508Q mutations and gRNA7 (5 and A10)-resistant modifications are shown in lower case letters. B)

Validation of the presence of E508Q mutation by Sanger sequencing after the cloning of E508Q KAT7 cDNA into the plasmid backbone. Top panel shows the targeted mutations at amino acid position 508 and the flanking nucleotides. Bottom panel shows the nucleotide readout from Sanger sequencing.

After generating the E508Q KAT7 mutant MOLM-13 cell line which expresses both E508Q KAT7 and the endogenous KAT7, we compared its growth (Figure 4.8A, grey line) with parental MOLM-13 without the expression of GFP (yellow line), MOLM-13 with GFP expression (blue line) and MOLM-13 expressing both wild-type KAT7 transgene and GFP (orange line). Compared to parental MOLM-13 cells with no GFP, expression of GFP protein reduces the growth of MOLM-13 cells marginally (Figure 4.8A). Additional expression of either wild-type KAT7 or E508Q KAT7 further reduces the

growth of MOLM-13 (Figure 4.8A). Simultaneous expression of GFP, E508Q KAT7 and endogenous KAT7 appears to reduce the growth of MOLM-13.

Next, we knocked-out the endogenous *KAT7* using gKAT7(A10) or gKAT7(5) to assess the effect of E508Q KAT7. We first evaluated if E508Q mutation gives rise to a catalytic dead KAT7, by western blot analysis. Whilst wild-type KAT7 rescues the knock-out phenotype, E508Q KAT7 cannot “write” H3K14ac and H4K12ac in MOLM-13 cells (Figure 4.8B), implying that this mutation abolishes the enzymatic activity of KAT7. Interestingly, we see reduced levels of H3K14ac and H4K12ac in cells expressing both E508Q and endogenous KAT7, although catalytic active KAT7 is present. This suggests that E508Q mutant has a dominant negative effect, by competing with endogenous KAT7 for histone substrates. In particular, we see H4K12ac level is more strongly affected compared to H3K14ac (Figure 4.8B). This dominant negative effect at a histone level did not induce short-term effects on the growth or proliferation of MOLM-13.

MOLM-13 harbouring E508Q phenocopied the KAT7 knock-out cells both in terms of their proliferation defect (Figure 4.8C) and the increase in CD11b expression (Figure 4.8D), suggesting that the catalytic activity of KAT7 is crucial for leukemic maintenance. We further demonstrated that the introduction of gKAT7(5) and gKAT7(A10)-resistant wild-type KAT7 transgene rescued the proliferation and differentiation phenotype, implying that the results from experiments using gKAT7 (5 and 10) is attributed to the on-target effects of the gRNAs used (Figure 4.8C-D).

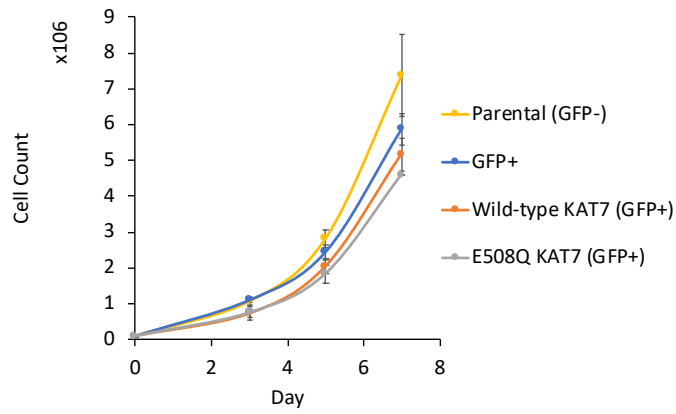
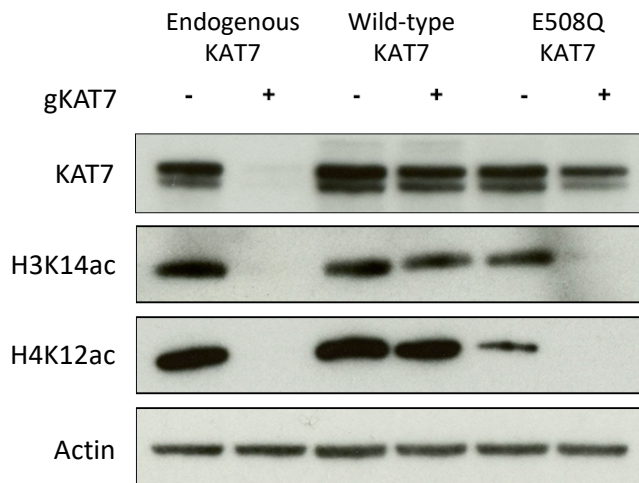
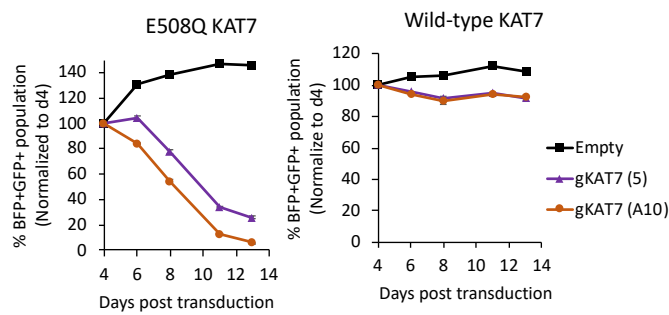
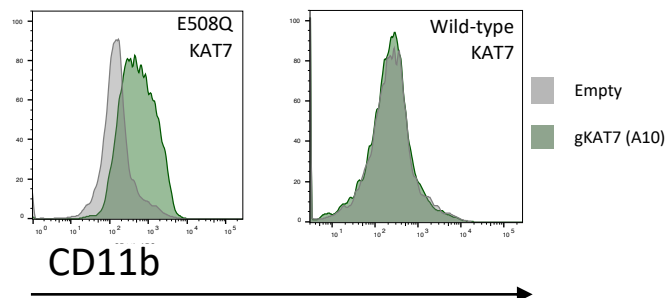
A**B****C****D**

Figure 4.8 E508Q catalytic dead mutant KAT7 affects the proliferation and differentiation of MOLM-13

A) Proliferation of parental MOLM-13 (GFP negative) and (GFP positive), MOLM-13 expressing wild-type KAT7 transgene (Wild-type KAT7) and MOLM-13 expressing E508Q mutant transgene (E508Q KAT7). B and D) MOLM-13 transduced with gKAT7 (A10) were sorted by flow cytometry on day 3 post-transduction to enrich for pure knock-out population. B) Changes in histone lysine acetylation levels after the knock-out of endogenous KAT7 in parental MOLM-13 cell line (Endogenous KAT7) and in MOLM-13 cells expressing either wild-type KAT7 transgene (wild-type KAT7) or E508Q mutant transgene (E508Q KAT7). C) Proliferation of MOLM-13 cells with endogenous KAT7 knock-out (BFP-positive) in the presence of E508Q mutant or wild-type KAT7 transgene (GFP-positive). Relative percentage of BFP and GFP double-positive cells were assayed overtime and all timepoints were normalized to day 4. Mean \pm S.D., n= 3. D) CD11b staining of MOLM-13 cells expressing either E508Q or wild-type transgene (green) and in MOLM-13 with endogenous KAT7 and the respective KAT7 transgene (grey), day 7 post-transduction.

4.5.3 Dominant negative effect of E508Q KAT7 on histone acetylation and cell proliferation

The dominant negative phenotype of E508Q prompted us to investigate further on the mutant's effect on H3K14ac and H4K12ac. Particularly, we are interested in whether one histone acetylation plays a more important role than the other in terms of maintaining leukemogenesis. One approach is to identify which KAT7 complex, KAT7-BRPF or KAT7-JADE is more "essential". We analysed the AML CRISPR-Cas9 screening dataset and gene expression profiles in the AML cell lines of all the members of KAT7 protein complex, including paralogues in the BRPF (BRPF1/2/3) family and JADE (JADE1/2/3) family (Figure 4.9). As mentioned in Chapter 1, KAT7-BRPF and KAT7-JADE has an affinity for specific lysine residues on H3 and H4 histone tails, respectively. Expect for JADE3, all members of the KAT7 complexes seems to be relatively highly expressed in the five AML cell lines (Figure 4.9). For MOLM-13 cell line, JADE2 and BRPF1 both showed similar levels of essentiality (\log_{10} (P-dropout)) as KAT7 (Figure 4.9). This data suggests that KAT7-JADE2 and KAT7-BRPF1 may be the predominant complexes in MOLM-13 that acetylates H4K12ac and H3K14ac, respectively. In MV4-11, BRPF1 and BRPF2 are more essential than JADE family (Figure 4.9), implying perhaps acetylation on H3 tail may be more important. JADE2 and BRPF3 are the strongest dropouts among the two families of proteins in OCI-AML2 (Figure 4.9), suggesting perhaps both H3K14ac and H4K12ac may be involved in the regulation of the pro-proliferation and anti-differentiation phenotypes. In HL-60 and OCI-AML3, the two cell lines in which KAT7 is not cell-essential, BRPF1

seems to be the most essential, although not as strongly depleted as in MOLM-13 and MV4-11 (Figure 4.9).

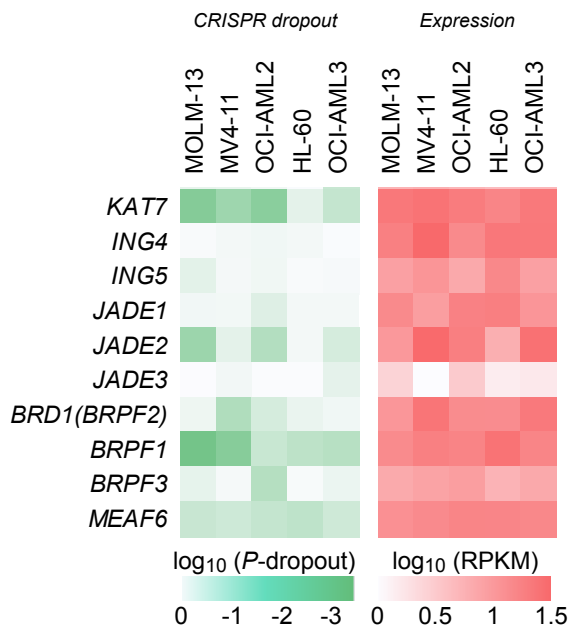


Figure 4.9 CRISPR-Cas9 depletion and gene expression of the members within the KAT7 protein complex

Left panel: Depletion p -value of each gene in the corresponding cell line. The lower the $\log_{10}(P\text{-dropout})$, the more significant the gene is considered a drop-out. Right panel: Gene expression in log-transformed reads per kilobase millions (RPKM) of each gene in the corresponding cell line. $\log_{10}(P\text{-dropout})$ values and transcriptomics data were obtained from the study performed Tzelepis *et al.*

To approach this question from another angle, we looked into the E508Q KAT7 mutant in MOLM-13 by sorting the bulk population into “low”, “mid” and “high” levels of mutant KAT7 expression via the levels of GFP signal. The aim is to explore the dominant negative effect of E508Q and its correlation with H3K14ac and H4K12ac levels as well as growth/proliferation. Notably, when sorting the MOLM-13 E508Q cells, we noticed two distinct GFP populations, which together constitute the bulk of the total cell population. We gated these two populations for cell sorting and named them “low” (brown) and “mid2” (purple) (Figure 4.10A). Additionally, we also sorted the population between “low” and “mid2” that we named population “mid1” (blue) as well as the populations at the extreme ends—“very low” (lime green) and “high” population (grey) (Figure 4.10A). The “very low”, “low”, “mid1”, “mid2” and “high” population constitutes 7.9%, 41.2%, 11.7%, 25.7% and 5.5% of the total population

respectively. Unfortunately, not enough cells were collected from the “very low” and “high” populations for western blot analysis but were sufficient for proliferation assay. Wild-type KAT7 MOLM-13 cells were also sorted by flow cytometry according to GFP expression, using the same gating as E508Q mutant.

Both H3K14ac and H4K12ac levels were significantly reduced in MOLM-13 despite having catalytic active KAT7 present in the cells (Figure 4.10B). In the varying concentration of E508Q KAT7 within the cell, higher “dose” of E508Q KAT7 protein, is associated with further reduction of H3K14ac and H4K12ac (compared to bulk), even with the presence of endogenous KAT7 (Figure 4.10B). This is particularly pronounced for H4K12ac as levels of this histone mark appears to be abolished in the “mid1” and “mid2” populations, whilst still detectable in the “low” population (Figure 4.10B). Conversely, H3K14ac has detectable levels across the “low”, “mid1” and “mid2” populations, although H3K14ac levels in the “mid2” population were significantly diminished compared to the bulk E508Q population (Figure 4.10B). In summary, cells with “low” amounts of E508Q has detectable levels of H3K14ac and H4K12ac, while “mid1” and “mid2” levels of this mutant KAT7 has detectable levels of H3K14ac but not H4K12ac. This recapitulates the findings in Figure 4.8B, where H4K12ac is relatively more affected than H3K14ac. In the wild-type KAT7 MOLM-13 population, higher levels of wild-type KAT7 transgene is associated with higher levels of H4K12ac, as seen with the “low”, “mid1” and “mid2” populations (Figure 4.10B). Regrettably, the exposed films of histone acetylation and KAT7 levels were not numerically quantified by densitometry for proper statistical analysis.

The proliferation of the GFP-sorted MOLM-13 populations with varying levels of KAT7 transgene was assessed over a 6-day time course, starting with 30000 cells in each well of a 96-well plate. At day 3, wild-type KAT7 populations appear to proliferate better than E508Q populations (Figure 4.10C). There is no apparent association between growth and the “dose” of mutant E508Q, among the sorted populations of E508Q KAT7 MOLM-13 (Figure 4.10C). MOLM-13 cells with wild-type transgene KAT7 has the highest cell number at day 3, further suggesting that having more KAT7 in the cell provide MOLM-13 with a growth advantage. Most populations except for E508Q “low”, reach at a comparable number of cells at day 6 (Figure 4.10C). Ideally, this assay should be conducted over a longer period, where the potentially varying growth rates between the different populations may become more pronounced. However, while maintaining the bulk E508Q MOLM-13 population over long term culture, we observed that these cells need to be repeatedly re-sorted to enrich for pure GFP-positive cells, as GFP-negative cells prevail over time, probably due to the growth disadvantage of expressing the E508Q mutant. It was not necessary to re-sort of GFP positive cells in wild-type MOLM-13, further supporting that KAT7 provides a growth advantage to cells.

Collectively, these data reflect that the level of H4K12ac is relatively more affected than H3K14ac by the presence of the E508Q mutant KAT7 (Figure 4.8B and Figure 4.10B). However, MOLM-13 E508Q can still proliferate, albeit at a slower rate, in the absence of H4K12ac (Figure 4.10B-C), meanwhile, these proliferating cells retain detectable levels of H3K14ac. These findings suggest that H3K14ac is necessary to sustain the leukemic programme in MOLM-13 cells, although H4K12ac may also play a role. In addition, the dominant negative effect of E508Q does not appear to have drastic consequences on the short-term proliferation but is unfavourable for MOLM-13 cells in long-term cell culture.

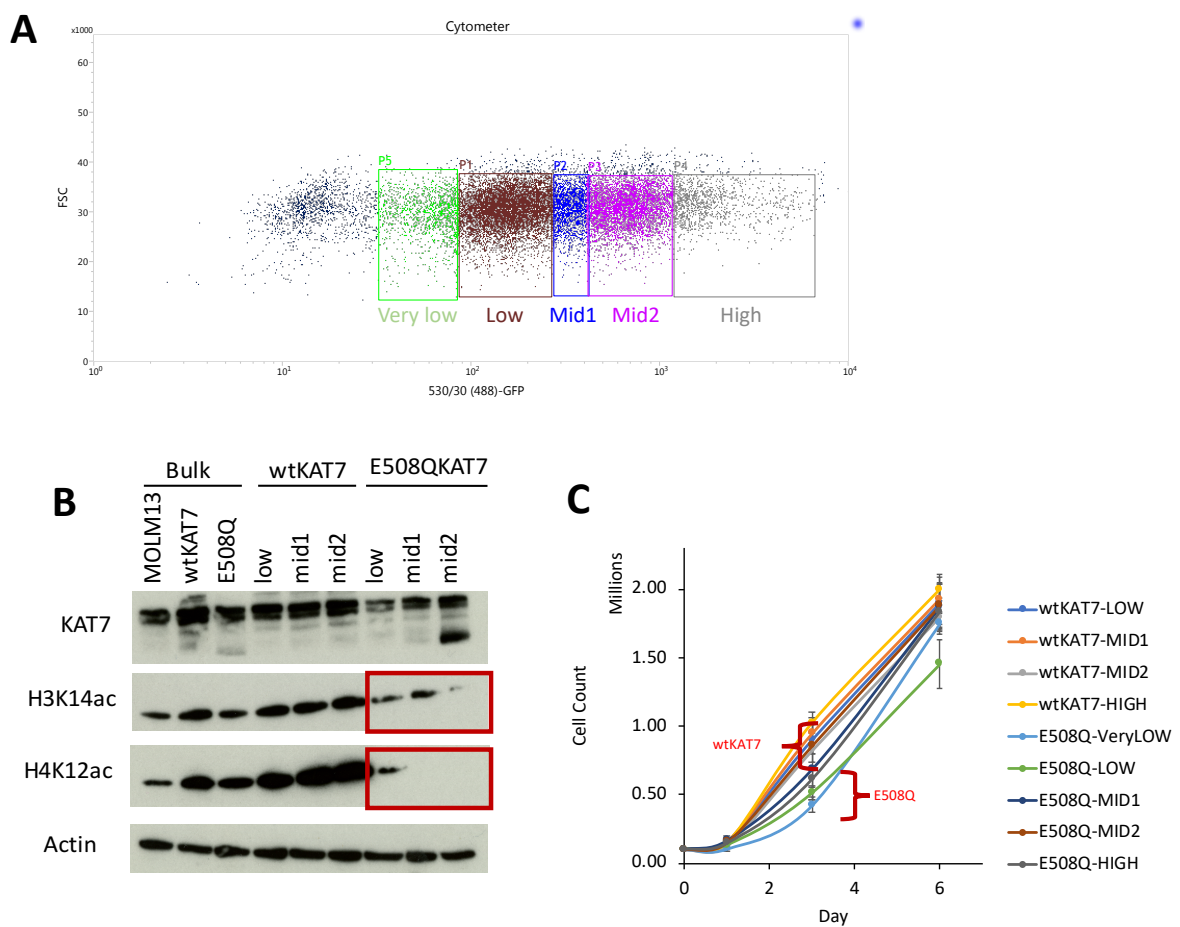


Figure 4.10 E508Q KAT7 mutant MOLM-13 populations with different levels of GFP expression

A) Flow-cytometry gating used for sorting E508Q KAT7 MOLM-13 cells into populations of different levels of GFP expression—very low, low, middle 1 (mid1), middle 2 (mid2), high. B) Immunoblot of the unsorted bulk population and the sorted wild-type KAT7 (wtKAT7) and E508Q KAT7 MOLM-13 populations with different levels of GFP/KAT7 transgene expression and their corresponding levels of H3K14ac and H4K12ac. The red box encircles the low levels of H3K14ac and H4K12ac in the sorted

population of E508Q KAT7 MOLM-13 cells, compared to the bulk. C) The proliferation of the wild-type KAT7 MOLM-13 and the E508Q KAT7 MOLM-13 cells with varying expression of GFP.

4.6 KAT7 predominantly occupies the gene promoter region

Having established that KAT7 is required for the global acetylation of H3K14 and H4K12, we next investigated the genomic location of KAT7 activity, using chromatin immunoprecipitation sequencing (ChIP-seq). We chose to sequence MOLM-13 and MV4-11 to represent cell lines in which KAT7 is essential and OCI-AML3 as the non-essential cell line. By examining the ChIP-seq data, we found a clear enrichment of KAT7 at promoters of protein-coding genes in MOLM-13, MV4-11 and OCI-AML3 cell lines (Figure 4.11A-B).

Promoter regions (± 2 kb from TSSs), despite constituting a small fraction of the genome, contain 63.3%, 40% and 48.5% of the KAT7 signals in MOLM-13, MV4-11 and OCI-AML3, respectively (Figure 4.11B). The total number of promoter peaks, by gene, in each cell line is indicated in Figure 4.11C. Notably, the number of genes bound by KAT7 at the promoter in MOLM-13 is approximately half of the promoter peaks found in MV4-11 and OCI-AML3 (Figure 4.11C). This observation could be explained, in part, by the immunoblot of the chromatin-immunoprecipitation (ChIP) (Figure 4.11D). In Figure 4.11D, the input of MOLM-13 showed two distinct KAT7 isoforms (~ 80 kDa and ~ 70 kDa), whilst only one isoform (~ 80 kDa) predominates in MV4-11 and OCI-AML3. The polyclonal KAT7 antibody used could pulldown the larger KAT7 isoform in MV4-11 and OCI-AML3, but not in MOLM-13. Furthermore, the ChIP in MOLM-13 pulls down the short isoform exclusively, and at a lesser amount than the ChIP in MV4-11 and OCI-AML3.

The immunoblot (Figure 4.11D) shows chromatin-immunoprecipitated IP (ChIP), whereby the KAT7 protein has been cross-linked with the chromatin using formaldehyde before immunoprecipitating with the KAT7 polyclonal antibody. This may suggest that in MOLM-13, only the short isoform of KAT7 binds the chromatin. Alternatively, the large isoform may also bind but if it forms a different protein complex, where the epitope recognized by the KAT7 antibody may be shielded within the macro-complex, which can also explain the relatively less efficient IP in MOLM-13 compared to the other two cell lines. It is not possible to distinguish between these two possibilities with the data presented.

The DNA element that has the next highest proportion of KAT7 occupancy is intergenic region and/or intron, depending on the cell line (Figure 4.11B). KAT7 has the lowest occupancy in the exon of protein-coding genes (Figure 4.11B). As KAT7 predominantly binds at gene promoter in all three AML cell lines, we subsequently focus on KAT7 promoter binding in the pathogenesis of AML at a chromatin level. Binding of histone acetyltransferases at the promoter and the subsequent acetylation of histones tails are typically associated with gene expression (Sheikh and Akhtar, 2018). Correlating KAT7 binding with gene expression, we found that the KAT7 occupancy at active promoters (purple) is significantly higher than inactive promoters (black) in MOLM-13 (Figure 4.11E-F), OCI-AML3 (Figure 4.11G-H) and MV4-11 (Figure 4.11I-J), which is consistent with the activating effect of promoter acetylation. Moreover, the highly active genes (red) has the highest KAT7 binding, reflecting a positive correlation between KAT7 promoter occupancy with gene expression in all cell lines.

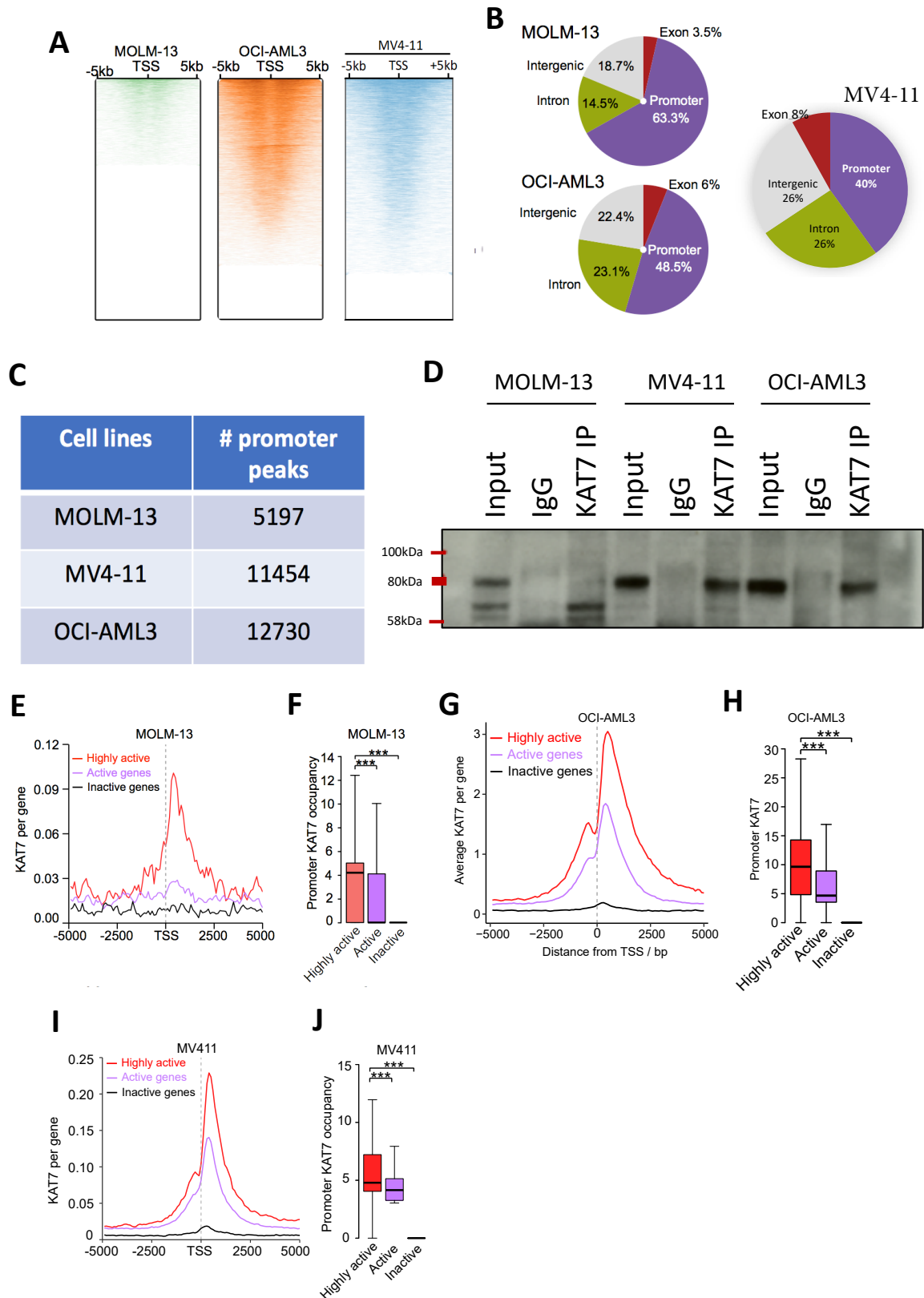


Figure 4.11 Genome-wide chromatin binding of KAT7 in MOLM-13, OCI-AML3 and MV4-11

A) KAT7 occupancy around 5kb windows at TSSs of protein-coding gene. B) Distribution of KAT7 occupancy at promoters (TSS±2kb), exons (excluding promoter exons), introns and intergenic regions quantified by MACS2 peaks. C) The total number of promoter binding peaks (by gene) from KAT7 ChIP-seq in each cell line. D) Immunoblot of KAT7 ChIP identical to the antibody and protocol used to generate KAT7 ChIP-seq. E, G and I) Average occupancy of KAT7 per gene per base of highly active genes (>10 FPKM), active genes (0-10 FPKM) and inactive genes (FPKM=0). F, H and J) Significant higher promoter (TSS±2kb) KAT7 in highly active genes than active genes and inactive genes by Mann-Whitney-Wilcoxon test. Highly active vs active: MOLM-13 ($p = 3.5 \times 10^{-227}$); OCI-AML3 ($p = 5.35 \times 10^{-273}$); MV4-11 ($p = 9.3 \times 10^{-100}$); highly active vs inactive: MOLM-13 ($p = 5.0 \times 10^{-324}$); OCI-AML3 ($p = 5.0 \times 10^{-324}$); MV4-11 ($p = 0$). Gene expression dataset was obtained from the study performed Tzelepis et al. ChIP-seq bioinformatics analysis and the corresponding figures are generated in collaboration with Muxin Gu.

4.7 Transcriptomics changes following KAT7 loss

To identify the genes which are regulated by KAT7, we performed RNA sequencing (RNA-seq) in MOLM-13, OCI-AML3 and MV4-11 for KAT7 knock-out cells and empty as control. From the time course experiment presented earlier in this chapter, KAT7 protein was reduced as early as day 3 post-transduction and was depleted by day 4. We, therefore, extracted mRNA from cells on day 3 and day 5 post-transduction to capture primary and secondary gene expression changes respectively.

Figure 4.12 showed the number of up and downregulated genes for MOLM-13, OCI-AML3 and MV4-11. Expectedly, day 3 has a significantly fewer number of differentially expressed genes compared to day 5 in all cell lines. By day 3 of KAT7 knock-out, we observed 317 up-regulated and 227 down-regulated genes in the MOLM-13 cell line. For MV4-11 and OCI-AML3, the transcriptomic response to KAT7 knock-out is slower and by day 5 reaches comparable numbers of differentially expressed genes as MOLM-13 day 3 (Figure 4.12). The relationship between KAT7 binding and gene expression changes in MOLM-13, MV4-11 and OCI-AML3, will be discussed further in the next chapter, particularly in the context of MLL-fusion transcriptional programme.

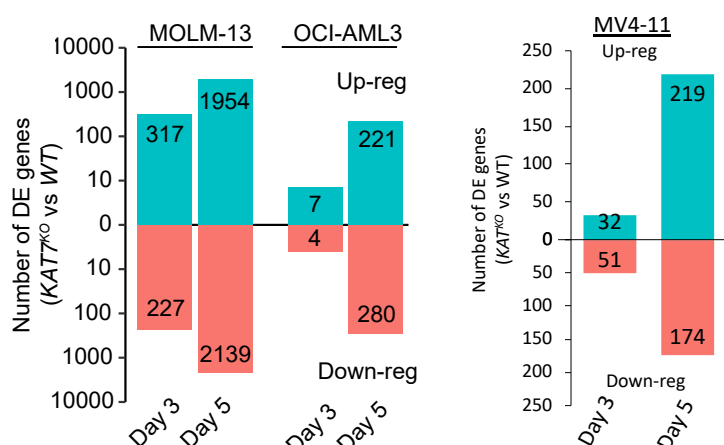


Figure 4.12 Gene expression changes after KAT7 knock-out

Number of differentially expressed genes in KAT7 knock-out versus empty control on day 3 and day 5 post-transduction in MOLM-13, OCI-AML3 and MV4-11. BFP-positive cells were sorted by flow cytometry on day 3 post-transduction and RNA were harvested at the respective time points. RNA-seq bioinformatics analysis and the corresponding figures are generated in collaboration with Muxin Gu and Swee Hoe Ong.

4.8 Discussion

The catalytic activity of KAT7 is essential for the proliferation of MLL-fusion AML. Drastic reductions in H3K14ac and H4K12ac can be seen following the loss of KAT7 protein and in E508Q catalytic dead KAT7, however, it remains unclear if one modification or a combination of acetylation at both residues are responsible for the phenotype. Our findings hint that H3K14ac might be more critical in MOLM-13, however, we cannot discount the additive effects of the lysine acetylation deposited by KAT7. Regrettably, we did not generate E508Q mutant in other AML cell lines therefore, unable to make broader conclusions on the dominant negative effects of the catalytic mutant of KAT7. Moreover, the extent to which non-histone protein acetylation by KAT7 in mediating transcription of target genes warrants further investigation as acetylation of nuclear and cytoplasmic protein such as transcription factors and tumour suppressors by other HATs have previously been reported (Roe et al., 2015; Sharma et al., 2000; Sheikh and Akhtar, 2018). Interestingly, both erythroblast and embryonic knock-down/knock-out of KAT7 in mice showed depletion of H3K14ac but not H4K12ac (Kueh et al., 2011; Mishima et al., 2011). KAT7 knock-down in erythroblasts also results in reduced H3K9ac levels, which

was not observed in our study (Mishima et al., 2011). Furthermore, H4K5ac and H3K9ac levels were found to increase after the knock-out of KAT7 in mouse embryos (Kueh et al., 2011). These observations further imply that the net change in acetylation upon KAT7 depletion is likely to be cell-type and context-specific. In contrast to the findings of Sauer et al., we did not observe significant changes in H4K5ac levels, at a global level.

Our attempt to generate H3K14ac ChIP-seq data has been unsuccessful, most likely due to the quality of the antibodies, as previously illustrated in Chapter 2. Possibly, H3K14ac levels may not be as abundant as other histone acetylation marks— which can also hinder the efficiency of ChIP experiments. Generating H3K14ac and H4K12ac ChIP-seq datasets from our cellular models will be pivotal in enhancing our understanding of whether these marks are important for the expression of KAT7-bound genes. Importantly, we confirmed that KAT7 promoter occupancy positively correlates with gene expression, typical of the effects of HATs and histone lysine acetylation (Sheikh and Akhtar, 2018).

Chapter 5: Mechanistic link between KAT7 and the pathogenesis of MLL-fusion AML

5.1 Introduction

Genetic ablation of KAT7 had the most pronounced anti-leukemic effects in AML cell lines harbouring the translocations involving the mixed-lineage leukemia (*MLL* or *KMT2A*) gene based on the findings described in Chapter 3. Partial tandem duplication (PTD) mutations of the *MLL* gene have also been found in patients, although less commonly than translocations of *MLL* (Ballabio and Milne, 2012; Meyer et al., 2013, 2018; Zeisig and So, 2015). Despite the fact that only 5-10% of adult acute leukemias possess the *MLL* gene fusion (Chen and Armstrong, 2015; Meyer et al., 2013, 2018; Slany, 2016), *MLL*-fusion leukemia generally characterize a subtype of the disease associated with a poor prognosis (Ayton and Cleary, 2001; Krivtsov and Armstrong, 2007; Papaemmanuil et al., 2016; Saultz and Garzon, 2016). The wild-type *MLL* gene encodes a histone methyltransferase that catalyzes the methylation of histone 3 lysine 4 residues (H3K4) and is crucial for the regulation of embryonic development and hematopoiesis (Jude et al., 2007; Yagi et al., 1998; Yokoyama, 2015, 2017; Yu et al., 1995). More than 70 fusion partners of *MLL* have been identified, however, the majority of the fusion partners are proteins involved in transcriptional elongation, namely AF9, AF4, AF6, AF10, ENL and ELL (Meyer et al., 2018; Slany, 2009, 2016; Zeisig and So, 2015). Interestingly, these fusion partners biochemically interact with one another and are part of the same protein complexes (Deshpande et al., 2012; Zeisig and So, 2015).

Target recognition and transcriptional activation of the *MLL*-fusion oncoprotein is hypothesized to be mediated through the *MLL* portion and the fusion partner portion, respectively (Yokoyama, 2015). *MLL*-fusion preferentially binds to non-methylated CpGs sites in the promoter-proximal regions via its CXXC domain of the N-terminal portion of the *MLL* retained in the fusion protein (Ayton and Cleary, 2001; Birke et al., 2002; Cierpicki et al., 2010; Yokoyama, 2017). In addition, target recognition of *MLL*-fusion protein is thought to depend on additional interactions of the N-terminal portion (*MLL*) with MENIN (Caslini et al., 2007; Grembecka et al., 2012; Kühn and Armstrong, 2015; Muntean and Hess, 2012; Yokoyama and Cleary, 2008; Yokoyama et al., 2005), LEDGF (Ashkar et al., 2018; Liedtke and Cleary, 2009; Okuda et al., 2014; Yokoyama and Cleary, 2008) and PAFc (Milne et al., 2010; Muntean and Hess, 2012; Muntean et al., 2010; Yokoyama, 2015, 2017). The C-terminal portion of the fusion protein, consisting of the fusion partner, can subsequently interact and form multiprotein complexes with the super elongation complex (SEC), AF4 family/ ENL family/ p-TEFb

(AEP) complex and DotCom complex (Ballabio and Milne, 2012, 2014; Deshpande et al., 2012; Zeisig and So, 2015).

p-TEFb (CDK9 and Cyclin T1/2) can be found in SEC, AEP and EAP complexes, whilst DOTCOM complex lacks p-TEFb but possesses DOT1L protein (Table 5.1). It is thus suggested that p-TEFb and DOT1L exist in a mutually exclusive manner with the MLL-fusion protein. However, some studies have reported the association of MLL fusion proteins with ENL associated protein (EAP) complex, which consists of both p-TEFb and DOT1L (Deshpande et al., 2012; Zeisig and So, 2015) (Table 5.1). The findings from these biochemical studies suggest that MLL-fusion does not act in isolation to regulate the leukemic transcriptional programme.

| Complex name | AEP | SEC | DotCom | EAP |
|-------------------|--------|--------|------------|---------------|
| Components | p-TEFb | p-TEFb | | p-TEFb |
| | AF4 | AF4 | | AF4 |
| | AF5Q31 | AF5Q31 | | AF5Q31 |
| | ENL | ENL | ENL | ENL |
| | | AF9 | AF9 | |
| | | | DOT1L | DOT1L |
| Others | | ELL1-3 | AF10, AF17 | LAF4, CBX8 |

Table 5.1 Multiprotein complexes associated with MLL-fusion proteins

In Chapter 4, we established a positive correlation between KAT7 promoter binding and gene expression. In this chapter, we explored the molecular mechanism by which KAT7 regulate the expression of MLL-fusion target genes.

5.2 Majority of the MLL-AF9 and MLL-AF4 spreading targets are bound by KAT7

Analysis of ChIP-seq in MOLM-13, MV4-11 and OCI-AML3 revealed that KAT7 predominantly occupied gene promoter, compared to other DNA regulatory elements, as described in Chapter 4. MLL-fusion protein binding to non-methylated CpG promoter-proximal regions of target genes is a popular model of gene activation by MLL-fusion (Kerry et al., 2017; Marschalek, 2011; Okuda et al.,

2014; Yokoyama, 2015, 2017). To find shared targets between KAT7 and MLL-fusion, we compared KAT7 promoter peaks with existing MLL-AF9 (from THP-1) and MLL-AF4 (from MV4-11) ChIP-seq datasets (Kerry et al., 2017), specifically those defined as “spreading targets”. Genes in which MLL-fusion bind to at the gene promoter and spread to the gene body are coined “spreading targets”, whilst genes that have promoter-restricted MLL-fusion binding patterns are termed “non-spreading targets” (Kerry et al., 2017). Furthermore, spreading targets are associated with higher gene expression compared to non-spreading targets and it is suggested that the expression of spreading MLL-fusion target genes are important for this subtype of leukemia (Kerry et al., 2017). The MLL-AF9 and MLL-AF4 spreading gene lists were applied filters to remove duplicates and non-coding genes, giving rise to 150 and 117 genes respectively. 35.5% (53 out of 150) of MLL-AF9 target genes are also bound by MLL-AF4, while 54.7% (53 out of 117) of MLL-AF4 spreading targets are bound by MLL-AF9 (Figure 5.1).

75.3% of MLL-AF9 spreading targets are bound by KAT7 at the promoter in MOLM-13 (Figure 5.2A). In the case of MV4-11, a striking 98.3% of MLL-AF4 spreading targets are bound by KAT7 at the promoter (Figure 5.2C). Majority of the MLL-AF9 and MLL-AF4 target genes are also bound by KAT7 in OCI-AML3, a cell line without MLL-fusion and this overlap is statistically significant (Figure 5.2B, D). As KAT7 additionally binds to promoter regions that are not occupied by spreading MLL-AF9 or MLL-AF4 and that OCI-AML3 (without MLL-fusion) has a large number of KAT7 peaks, together suggest that the genome-wide chromatin occupancy of KAT7 may not be dependent on MLL-fusion.

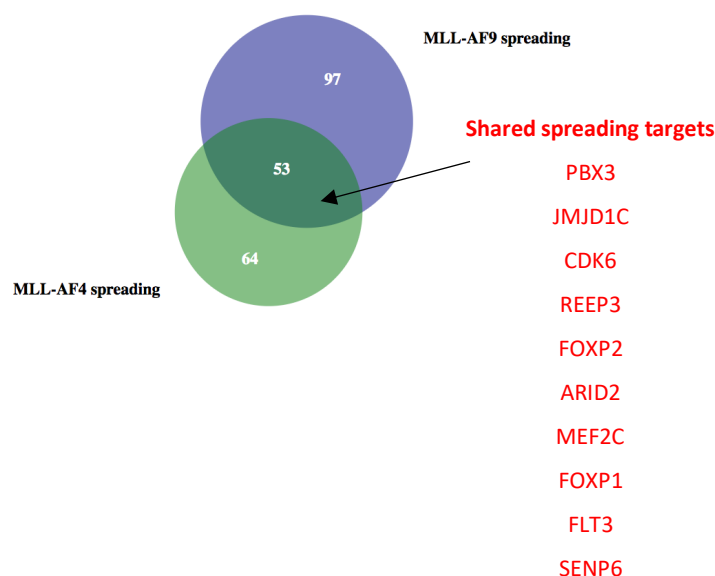


Figure 5.1 Shared spreading targets of MLL-AF9 and MLL-AF4

MLL-AF9 and MLL-AF4 ChIP-seq dataset from THP-1 and MV4-11 respectively (Kerry et al., 2017). Spreading defines the binding pattern of MLL-fusion at the gene body in addition to the promoter of the target genes. The number of unique, protein-coding targets for MLL-AF9 ($n = 150$) and MLL-AF4 ($n = 117$). Notable MLL-fusion spreading targets that are shared between MLL-AF4 and MLL-AF9 are listed.

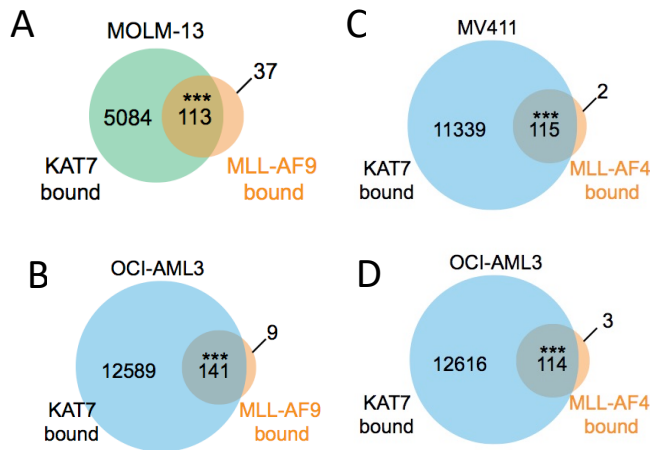


Figure 5.2 Majority of MLL-AF9 and MLL-AF4 spreading targets are co-bound by KAT7 at the promoter

Significant overlap between KAT7-bound and MLL-AF9 spreading targets in A) MOLM-13 ($p = 3.9 \times 10^{-34}$) and B) OCI-AML3 ($p = 4.8 \times 10^{-15}$) and between KAT7-bound and MLL-AF4 spreading targets in C) MV4-11 ($p = 1.4 \times 10^{-22}$) and D) OCI-AML3 ($p = 3.6 \times 10^{-16}$) by Fisher's Exact test.

5.3 Top KAT7 bound genes are MLL-fusion targets in MOLM-13 and MV4-11 but not in OCI-AML3

To investigate the relationship between KAT7 binding and MLL-fusion binding, we ranked KAT7 occupied promoter from low-to-high in MOLM-13, MV4-11 and OCI-AML3. Strikingly, we observed that the top KAT7-bound genes in MOLM-13 are MLL-AF9 exclusive spreading targets or are shared targets between MLL-AF9 and MLL-AF4 (Figure 5.3A). Correspondingly, genes with the highest KAT7 promoter occupancy in MV4-11 are either MLL-AF4 exclusive spreading targets or targets co-bound by MLL-AF9 and MLL-AF4 (Figure 5.3B). In contrast, genes with the highest KAT7 promoter occupancy in OCI-AML3 are neither MLL-AF9 or MLL-AF4 bound genes (Figure 5.3C-D). This may suggest that KAT7 recruitment to these specific promoter sites may be MLL-fusion dependent in the context in which MLL-AF9 or MLL-AF4 is present.

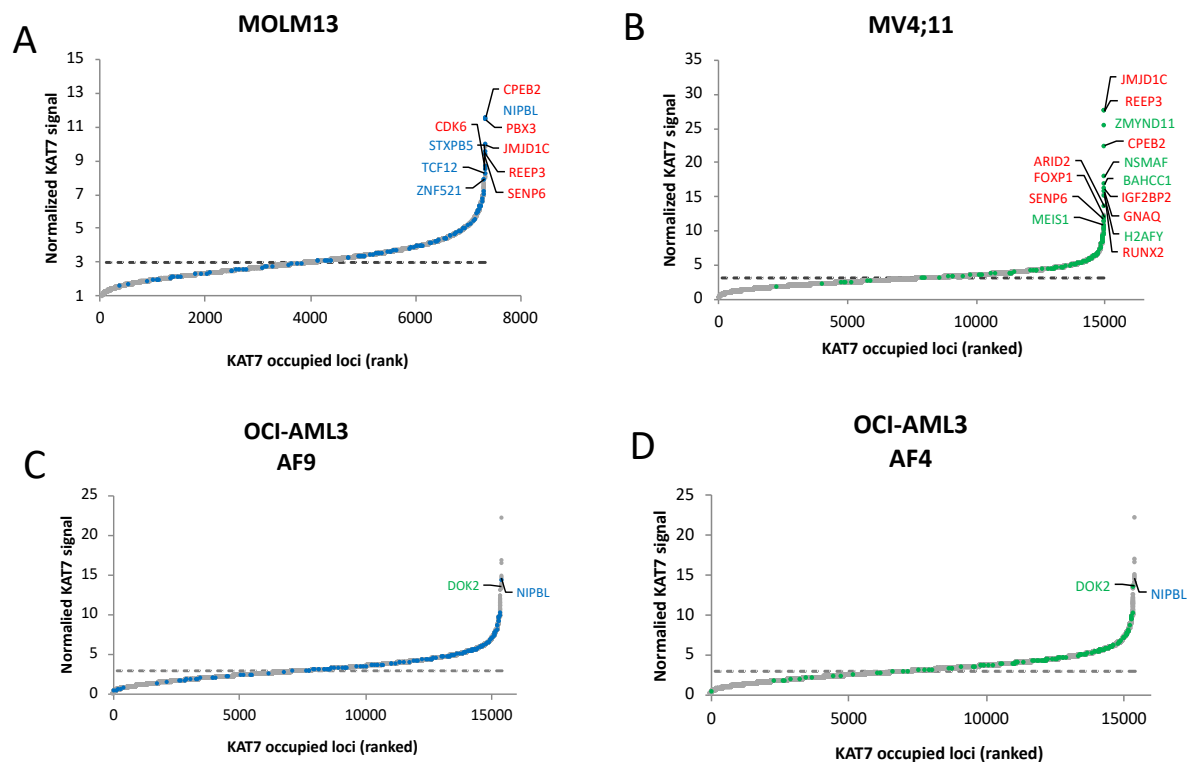


Figure 5.3 Ranked normalized KAT7 promoter occupancy signal

KAT7 promoter occupancy was normalized for comparison of ChIP-seq signal between cell lines. Blue and green dots represent the MLL-AF9 (THP-1) and MLL-AF4 (MV4-11) spreading genes respectively, dataset derived from Kerry et al. Top KAT7-bound genes in A) MOLM-13 and B) MV4-11 are bound by MLL-AF9 (blue), MLL-AF4 (green) or shared targets of MLL-AF9 and MLL-AF4 (red). C-D) Top KAT7-bound genes in OCI-AML3 are not targets of MLL-AF9/MLL-AF4. Normalisation was based on the assumption that promoter occupancy of KAT7 is unchanged for the majority of protein-coding genes. Promoter occupancies (± 1000 bp of TSS) of all the 20291 protein-coding genes were calculated for each background-subtracted sample and normalised to the median value of promoter occupancies.

5.4 Genes co-bound by MLL-AF9 and KAT7 at promoter are downregulated after KAT7 loss in MOLM-13

We next investigated whether KAT7 regulates the transcription of MLL-AF9 targets by examining how gene expression changes in KAT7 knock-out cells compared to wild-type using RNA-seq. The up and down-regulation of gene expression may be directly affected by the KAT7 loss or due to indirect effects. To establish which genes are directly regulated by KAT7, we separated the genes that are occupied by KAT7 at the promoter with those whose promoters are not bound by KAT7 and examined whether

their expression levels change following *KAT7* knock-out. We found that *KAT7*-bound genes that demonstrate spreading MLL-AF9 patterns in MOLM-13 are almost exclusively associated with decreased transcript levels at both day 3 and day 5 (Figure 5.4A) in marked contrast to the same genes in the OCI-AML3 cell line (Figure 5.4B). In MOLM-13 day 3, 20 (of 150) MLL-AF9 spreading targets are significantly downregulated (Figure 5.4C) and these genes display significantly higher *KAT7* occupancy than upregulated (n=4) or non-differentially expressed (n=126) genes (Figure 5.4D). These findings are in keeping with down-regulation likely caused by the *KAT7* loss at promoters, whereas the up-regulation is likely caused by indirect effects.

Notably, no association between high *KAT7* occupancy and down-regulation was observed in OCI-AML3, a cell line which lacks an MLL-fusion oncoprotein (Figure 5.4E-F). Furthermore, in MOLM-13 only genes that are bound by both *KAT7* and demonstrate an MLL-AF9 “spreading” pattern are down-regulated, whilst the ones bound only by *KAT7* are not (Figure 5.4G). These data suggest that *KAT7* likely operates in conjunction with MLL-AF9 to activate leukemogenic genes at these specific loci. As expected, the OCI-AML3 cell line displays no significant difference in expression changes with respect to the binding of *KAT7* at a genome-wide level or among MLL-AF9 bound targets (Figure 5.4H).

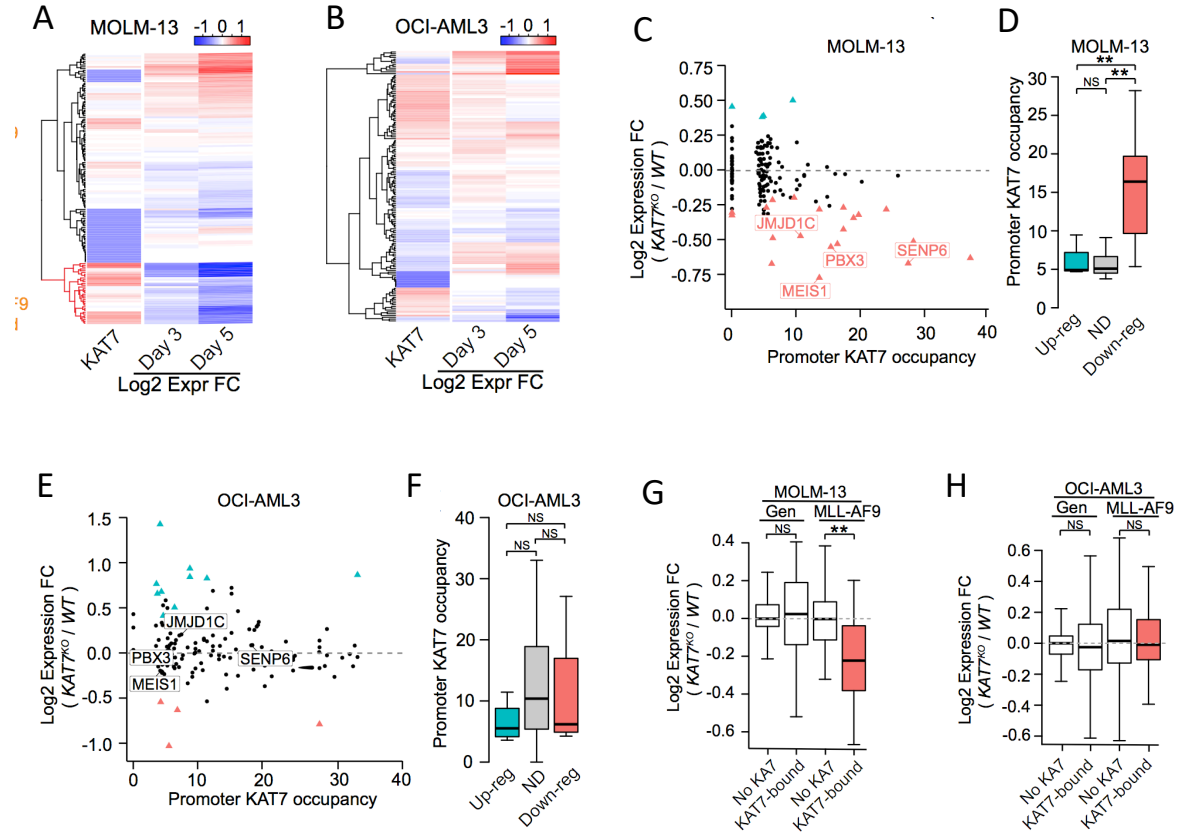


Figure 5.4 Transcriptomics change after KAT7 knock-out and the chromatin occupancy of KAT7 in MOLM-13 and OCI-AML3

A-B) Heat map showing promoter KAT7 occupancy (Z-transformed \log_{10} values), \log_2 fold changes of expression between KAT7 knock-out and wild-type (WT) in day 3 and day 5, for the 150 MLL-AF9 spreading genes. Data for heat maps were unscaled and ordered by hierarchical clustering of Euclidian distances. C) Scatter plot of promoter KAT7 occupancy against the change of expression in KAT7 knock-out for MLL-AF9 targets and MEIS1. Significantly up- and down-regulated genes (adjusted $p < 0.05$ by DESeq2) are highlighted in blue and red respectively. D) Significantly higher promoter KAT7 in down-regulated genes compared to up-regulated genes ($p = 0.0027$) and genes that are not differentially expressed (ND, $p = 2.7 \times 10^{-7}$) by one-tailed T-test. ND versus up-reg $p = 0.96$. E) Same as C for OCI-AML3. F) Same as D for OCI-AML3 showing no significant difference. Down-reg versus up-reg $p = 0.75$, down-reg versus ND $p = 0.59$ and up-reg versus ND $p = 0.08$. G) Distribution of change of gene expression (\log_2 scale) for KAT7-bound (>10 normalized units) and unbound (zero occupancy) genes in MOLM-13. On the genomic scale (Gen) two-tailed T-test $p = 0.51$ (18,622 unbound versus 85 bound genes) and within MLL-AF9 targets $p = 4.3 \times 10^{-5}$ (123 unbound versus 27 bound genes). H) Same as G for OCI-AML3. Genomic $p = 0.057$ (14766 unbound, 3941 bound) and MLL-AF9 targets $p = 0.54$ (78 unbound, 72 bound).

5.5 Loss of KAT7 results in general gene expression changes in MV4-11 that is not associated with MLL-AF4 spreading targets

Binding of KAT7 does not show an obvious correlation with the gene-expression change of MLL-AF4 spreading targets in MV4-11 on day 3 and day 5 after KAT7 knock-out (Figure 5.5A). Unexpectedly, the changes in gene expression on day 3 and day 5 appears to be contrasting (Figure 5.5A). Majority of the targets co-bound by spreading MLL-AF4 and KAT7 are not differentially expressed between KAT7 knock-out and empty control on day 3 post-transduction in MV4-11 (Figure 5.5B). On day 3, only 9 (of 117) MLL-AF4 targets are significantly downregulated, and the remaining (n=108) are not differentially expressed (Figure 5.5B-C). The targets co-bound by MLL-AF4 and KAT7 targets which are significantly downregulated on day 3 are *NUDT5*, *JMJD1C*, *PTEN*, *FLT3*, *NUSAP1*, *RAB27A*, *MRPL33*, *ITGA4* and *SATB1*. Notably, there is no significant difference in KAT7 promoter occupancy between down-regulated and non-differential expressed MLL-AF4 spreading target genes (Figure 5.5C), contrasting the trends observed in MOLM-13 (Figure 5.4C-D).

When comparing MLL-AF4 targets in terms of those that are bound or unbound by KAT7 at the promoter, those that are occupied by KAT7 (n=60) are downregulated compared to loci that are unbound by KAT7 (n=57) on day 3 and the difference in gene expression between these two population of genes is statistically significant (Figure 5.5D). This is also the trend at a genome-wide level, excluding MLL-AF4 spreading targets (Figure 5.5D). These data suggest that the loss of KAT7 in MV4-11 results in the general downregulation of genes bound by KAT7 at the promoter and only a minor subset of MLL-AF4 spreading target genes are significantly downregulated despite the majority of the MLL-AF4 spreading genes are mutual targets of KAT7.

Unexpectedly, although being co-bound by KAT7 and MLL-AF4, established MLL-fusion targets such as *MEIS1* and *PBX3* did not appear to be differentially expressed in KAT7 knock-out compared to empty control by RNA-seq on day 3 and day 5 post-transduction. Therefore, we analysed a set of well-studied MLL-fusion targets, most of which are associated with spreading MLL-AF4 (*PBX3*, *MEIS1*, *CDK6*, *BCL2*, *JMJD1C* and *SENK6*) and some associated with non-spreading MLL-AF4 (*HOXA9* and *MYC*). We also included a few genes of the *HOXA* cluster that are not bound by MLL-AF4 in MV4-11 (*HOXA5*, *HOXA6*, *HOXA10* and *HOXA11*) to investigate how they change in gene expression. We analysed the same RNA extracts that were submitted for RNA-seq by RT-qPCR to compare the output of two different methods that assess transcriptomics changes. Spreading MLL-AF4 targets *PBX3*, *MEIS1*, *CDK6*, *BCL2*, *JMJD1C* and *SENK6* are strongly downregulated in KAT7 knock-out compared to wild-type by RT-qPCR on day 5 (Figure 5.5E). Only *JMJD1C* was significantly differentially

expressed in RNA-seq analyses. Genes not bound by MLL-AF4 in MV4-11 such as *HOXA10* and *HOXA11* also appears to be downregulated by RT-qPCR. The inconsistency between RNA-seq and RT-qPCR results should be further investigated.

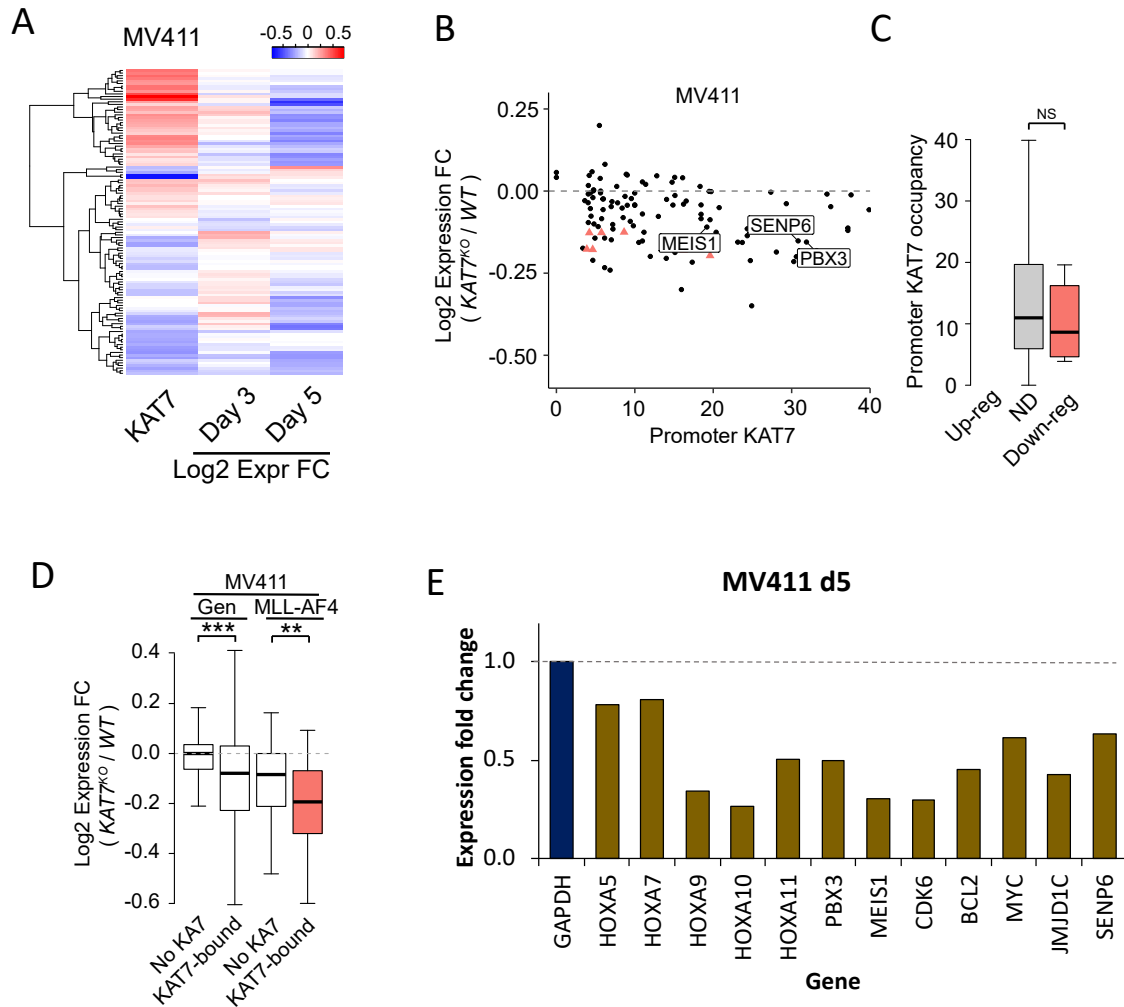


Figure 5.5 Transcriptomics changes in MV4-11 after KAT7 knock-out and the chromatin occupancy of KAT7

A) Heat map showing promoter KAT7 occupancy (Z-transformed \log_{10} values), \log_2 fold changes of expression between KAT7 knock-out and wild type (WT) in day 3 and day 5, for the 117 MLL-AF4 spreading genes. Data for heat maps were unscaled and ordered by hierarchical clustering of Euclidian distances. B) Scatter plot of promoter KAT7 occupancy against the change of expression in KAT7 knock-out for MLL-AF4 targets. Significantly up- and down-regulated genes (adjusted $p < 0.05$ by DESeq2) are highlighted in blue and red respectively. C) No significant difference in KAT7 occupancy in down-regulated genes compared to genes that are not differentially expressed genes (ND, $p = 0.39$) by one-

tailed T-test. D) Distribution of change of gene expression (\log_2 scale) for KAT7-bound (>10 normalized units) and unbound (zero occupancy) genes in MV4-11. On the genomic (Gen) scale two-tailed T-test $p = 2.7 \times 10^{-42}$ (17619 unbound versus 1121 bound genes) and within MLL-AF4 targets $p = 2.2 \times 10^{-3}$ (57 unbound versus 60 bound genes). E) Changes in gene expression of established MLL-fusion targets between KAT7 knock-out and empty control by RT-qPCR using the same RNA extract submitted for RNA-seq.

5.6 Auxin-inducible degron (AID) system for rapid KAT7 protein degradation

To study the acute effects of KAT7 depletion, we utilized the auxin-inducible degron (AID) system to rapidly eliminate KAT7 protein in cells (Nishimura et al., 2009). This system requires two components. Firstly, the protein of interest needs to be tagged with a mini-AID (mAID) peptide, a 68 amino acid fragment of the original AID (Natsume et al., 2016; Nishimura and Kanemaki, 2014). Secondly, the plant-specific F-box protein transport inhibitor response 1 (*TIR1*) gene, needs to be introduced as it is not expressed in mammalian system (Natsume et al., 2016; Nishimura and Kanemaki, 2014). Upon the treatment of plant hormone auxin, the protein of interest is poly-ubiquitinated by E3 ubiquitin ligase and rapidly degraded by proteasome within the cell (Natsume et al., 2016; Nishimura and Kanemaki, 2014; Nishimura et al., 2009).

Briefly, we generated stable *KAT7-mAID* expressing MOLM-13 and MV4-11 cells followed by knocked out endogenous KAT7 with gKAT7 (A10), such that the only KAT7 protein in cells is the tagged KAT7-mAID, before introducing the *OsTIR1* gene (*TIR1* gene derived from *Oryza sativa*). The MOLM-13 AID and MV4-11 AID cells were generated after successive rounds of fluorescent marker selection (Figure 5.6A). We first confirmed that the proliferation of both MOLM-13 AID and MV4-11 AID cells were similar to that of the parental (Figure 5.6B), indicating that the expression of mAID-tagged KAT7 and ectopic expression of *OsTIR1* was not detrimental to the cells.

Then, using this system, we performed time-course experiments following treatment of indole-3-acetic acid (IAA), which is the most abundant hormone of the auxin class (Lavy and Estelle, 2016). A dramatic reduction of KAT7 protein levels is observed as early as 2 hours after treatment with IAA in both MOLM-13 AID (Figure 5.6C) and MV4-11 AID cells (Figure 5.6D). This was associated with the complete loss of H3K14ac over the same time course (Figure 5.6C-D). Notably, IAA needs to be replenished by 72 hours, as KAT7 levels in MV4-11 is restored following 3 consecutive days of IAA treatment (Figure 5.6D). Furthermore, the addition of IAA to parental MOLM-13 and MV4-11 cells do

not affect the levels of KAT7 protein or H3K14ac (Figure 5.6C-D), thus reflect that levels of endogenous proteins, particularly KAT7, without mAID-tagged is unaltered by the addition of this plant hormone.

Two isoforms of KAT7 can be observed in the immunoblot of MOLM-13 AID and MV4-11 AID cells, (Figure 5.6C-D), potentially due to post-translational modifications on KAT7 protein that alter its molecular weight. Noticeably, parental MOLM-13 also has two distinct isoforms of KAT7, but this is not as distinct in the parental MV4-11. It is unknown if the different isoforms have different roles in the pathogenesis of AML.

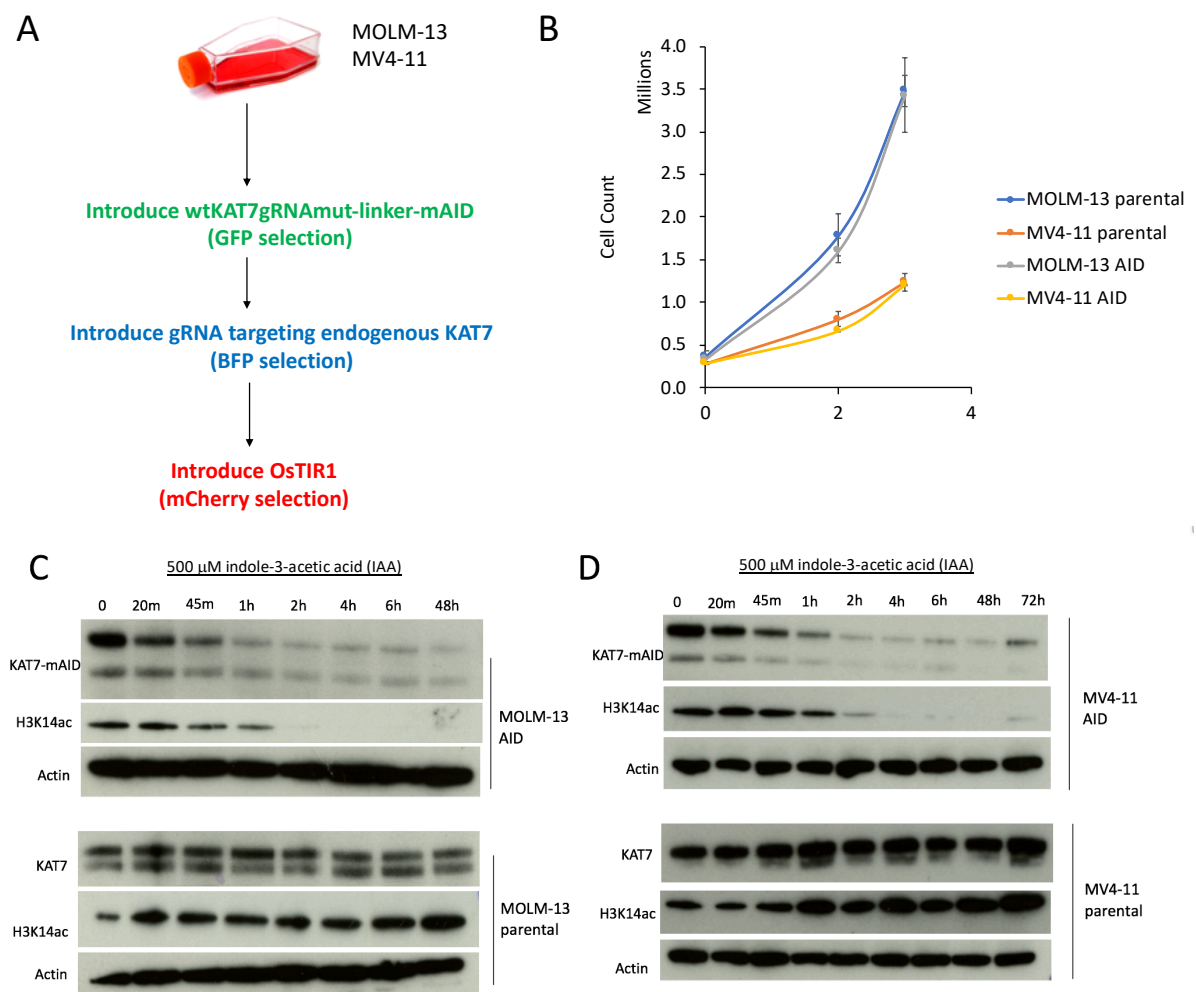


Figure 5.6 Auxin-inducible degron (AID) MOLM-13 and MV4-11 induces rapid KAT7 protein degradation and H3K14ac depletion

A) Schematic illustrating the sequential rounds of transduction and selection to generate the AID MOLM-13 and MV4-11 cell lines. B) Cell proliferation of parental and AID MOLM-13 and MV4-11 cells. C) Time-course experiment of indole-3-acetic acid (IAA) treatment and the corresponding levels of

KAT7 protein and H3K14ac in MOLM-13 AID cells (top panel) and parental MOLM-13 cells (bottom panel). D) Same as C for MV4-11. m, minutes; h, hours.

5.7 Gene expression changes and the differentiation of IAA treated MOLM-13 AID and MV4-11 AID cells

To address if KAT7 knock-out AID cells recapitulate gRNA-mediated KAT7 knock out cells, we assessed the differentiation phenotype and changes in gene expression of some of the classical MLL-fusion targets such as *PBX3*, *MEIS1*, *CDK6* and *BCL2* (Dawson et al., 2011; Guenther et al., 2008; Kohlmann et al., 2005; Krivtsov et al., 2006; Liedtke and Cleary, 2009; Neff and Armstrong, 2013) following treatment of IAA in MOLM-13 AID and MV4-11 AID cells. These MLL-fusion target genes are all downregulated, albeit at different rates in the different cell lines (Figure 5.7A-B). For instance, *SENK6*, *PBX3* and *JMJD1C*, which are bound by KAT7 in MOLM-13 and MV4-11 and are spreading target of both MLL-AF9 and MLL-AF4, is strongly reduced in MOLM-13 but to a lesser extent in MV4-11 over a 48-hour time course. Furthermore, some genes such as *MEIS1* and *JMJD1C* are significantly downregulated by 2 hours in MOLM-13. In general, the decrease in gene expression in MV4-11 AID cells is less pronounced compared to MOLM-13 AID cells over the same time course. Furthermore, myeloid differentiation (CD11b) is observed after 48 hours of IAA treatment in MOLM-13 AID, mirroring the phenotype seen CRISPR-Cas9 mediated KAT7 knock-out. A 41.1% increase in CD11b expression at 48 hours of IAA treated MOLM-13 AID compared to cells without IAA treatment (cells treated with water as control) (Figure 5.7C), corresponds to the differentiation phenotype on day 5 of gRNA-mediated KAT7 knock-out. This suggests that the auxin-inducible system leads to a more acute effect compared to the CRISPR-Cas9 system, therefore the former is a suitable system to assess primary changes immediately following the loss of KAT7 protein. No CD11b differentiation is observed in IAA treated MV4-11 AID cells compared to water-treated control (Figure 5.7D). This reinforces the lack of differentiation seen in the gKAT7-mediated knock-out MV4-11 cells in Chapter 3.6.3, strengthening the evidence that an increase in CD11b differentiation is not a consequent of KAT7 loss in this cell line.

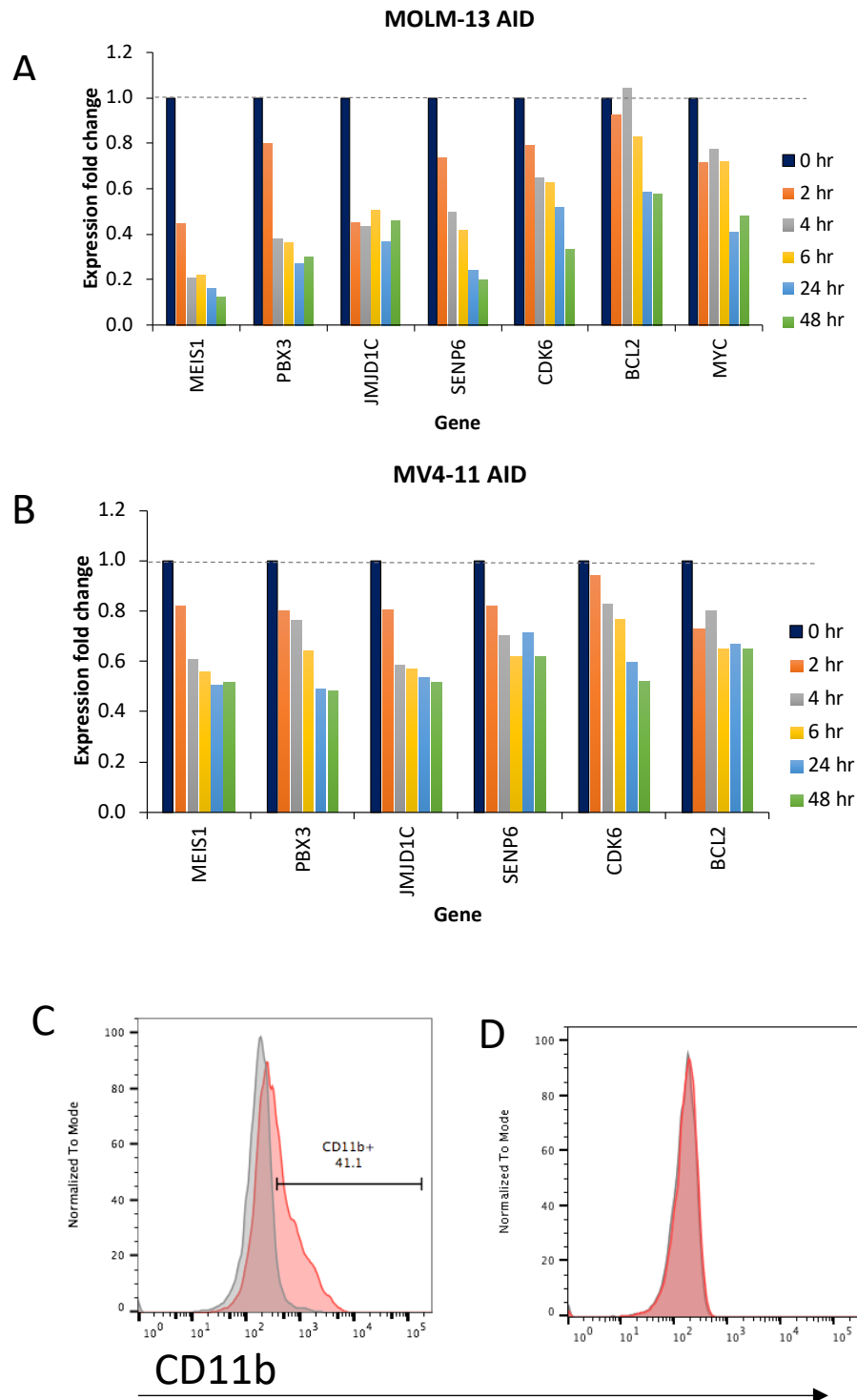


Figure 5.7 Changes in gene expression and differentiation after IAA-induced KAT7 depletion in AID MOLM-13 and MV4-11 cells

A-B) Treatment of IAA and the downregulation of well-known MLL-fusion targets over a 48-hour time course in A) MOLM-13 AID and B) MV4-11 AID cells. C) Increase in CD11b expression in IAA treated MOLM-13 AID (red histogram) compared to H₂O treated control (grey histogram) at 48 hours of IAA

treatment. D) Similar levels of CD11b expression in IAA treated and H₂O treated MV4-11 AID cells at 48 hours of IAA treatment.

5.8 Loss of KAT7 and the occupancy of MLL-fusion proteins

Given that the majority of MLL-AF4 and MLL-AF9 spreading targets are bound by KAT7, we sought to examine if KAT7 is needed for the recruitment of MLL-fusion to its target at the promoter region. Table 5.2 summarizes the genes of interest and whether its promoter region is bound by spreading (SP) or non-spreading (NSP) MLL-AF9 and MLL-AF4. The time point at which the genes are significantly downregulated in the KAT7 knock-out RNA-seq dataset for MOLM-13 and MV4-11 are also indicated.

| Gene | KAT7 | MLL-AF9 | MLL-AF4 | MOLM-13 (day) | MV4-11 (day) |
|---------------|------|---------|---------|---------------|--------------|
| <i>JMJD1C</i> | ✓ | SP | SP | 3 | 5 |
| <i>SENP6</i> | ✓ | SP | SP | 3 | |
| <i>HOXA9</i> | ✓ | NSP | NSP | | |
| <i>HOXA10</i> | ✓ | NSP | | 3 | |
| <i>PBX3</i> | ✓ | SP | SP | 3 | |
| <i>MEIS1</i> | ✓ | | SP | 3 | |
| <i>BCL2</i> | ✓ | NSP | SP | 5 | |

Table 5.2 Selective established MLL-fusion targets and the binding profiles of KAT7, MLL-AF9, MLL-AF4 and the earliest time point of gene down-regulation following KAT7 knock-out.

Selective KAT7 targets that are also well-known targets of MLL-fusions. A tick in the KAT7 column means there is binding of KAT7 at the promoter region of the gene in both MOLM-13 and MV4-11. MLL-AF9 and/or MLL-AF4 occupancy with spreading (SP) or non-spreading (NSP) patterns at the respective genes are indicated. The earliest timepoint (day post-transduction) at which significant gene expression changes are detected in the respective cell lines after KAT7 knock-out from RNA-seq analysis is indicated in the last two columns. All gene expression changes indicated in the table are associated with downregulation.

Chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR) was performed in MOLM-13 AID and MV4-11 AID cells. The MLL N-terminal, AF9 C-terminal and AF4 C-terminal antibodies were used to probe for the chromatin occupancy of MLL-AF9 and MLL-AF4 binding in

MOLM-13 and MV4-11. In addition to the promoter of the shared targets of MLL-AF9/MLL-AF4 and KAT7, we also assayed the KAT-bound introns and enhancer regions of these genes.

5.8.1 MLL-AF9

In MOLM-13, treatment with IAA for 24 hours leads to an insignificant increase in occupancy of MLL-N and AF9-C at the promoter of *JMJD1C* and *SENK6* (Figure 5.8A and C). A marginal increase in binding of MLL-AF9 upon the loss of KAT7 is similarly observed for *HOXA9* and *HOXA10*, although more replicates are needed to determine statistical significance (Figure 5.8A, C). For the promoter and intron binding at *PBX3*, a slight increase can be seen in the MLL-N portion but the contrary is observed for AF9-C terminal (Figure 5.8A, C). This inconsistency could reflect noise or the difference in occupancy of the wild-type MLL and wild-type AF9 at the *PBX3* loci. Although MLL-AF9 ChIP-seq data derived from THP-1 did not show binding at *MEIS1* (Kerry et al., 2017; Prange et al., 2017), our ChIP-qPCR suggests that MLL-AF9 might bind to *MEIS1* at both promoter and intron regions in MOLM-13. However, we cannot eliminate that the ChIP-qPCR signal is picking up the wild-type MLL and wild-type AF9 binding, therefore needs to be verified by performing MLL-C and AF9-N ChIP-qPCR. Moreover, the binding of MLL-N and AF9-C at *BCL2* promoter and enhancer region is relatively weak and there are minimal changes in occupancy at these loci after the loss of KAT7 (Figure 5.8A, C).

As a control, we performed AF9-C terminal ChIP-qPCR for the same targets in MV4-11 (Figure 5.8D), and found low levels of AF9 occupancy at these loci, compared to MOLM-13. This is reflected by the low ChIP signal levels (percentage input) across all the loci and these are also similar to the binding signals detected at a negative, gene desert region (Figure 5.8D). Furthermore, the low percentage of input is unlikely to be attributed by poor antibody quality, as identical experimental set up using the same AF9 antibody in MOLM-13 yield approximately 100-times stronger signal (Figure 5.8C, D). Assuming these loci are important for leukemogenesis in MV4-11, this finding could suggest that the MLL-AF4 associated protein complex recruited to these loci may not include the wild-type AF9.

5.8.2 MLL-AF4

After 24 hours of treatment with IAA, MLL-AF4 show similar trends of increase at all of the above-mentioned MLL-fusion target loci in MV4-11, except for *PBX3* intron (Figure 5.8B, F). Notably, signals of MLL-N and AF4-C at all of these loci (Figure 5.8B, F) are significantly stronger than AF9-C

in MV4-11 (Figure 5.8D). Repeats are needed to determine the statistical significance. Strikingly, we observed a statistically significant decrease in wild-type AF4 binding at these same target loci in MOLM-13 (Figure 5.8E). This suggests that KAT7 loss has an effect on the binding of AF4 in MOLM-13, which could further imply that the MLL-AF9 macromolecule complex in MOLM-13 consists of AF4.

Notably, ChIP-qPCR of MLL-N, AF9-C and AF4-C will detect signals from the wild-type counterparts as well as the MLL-fusion protein. Therefore, to distinguish between fusion vs wild-type, we need to assay MLL-C and AF9-N/AF4-N to account for the signals from the wild-type. If the slight increase in the recruitment of MLL-AF9 and MLL-AF4 observed at some loci following the loss of KAT7 is true, this could reflect some sort of compensation mechanism in order to restore the expression of these genes, at least in MOLM-13. However, the data thus far detects no significant changes in MLL-AF9 and MLL-AF4 promoter occupancy at the target loci in MOLM-13 and MV4-11, respectively.

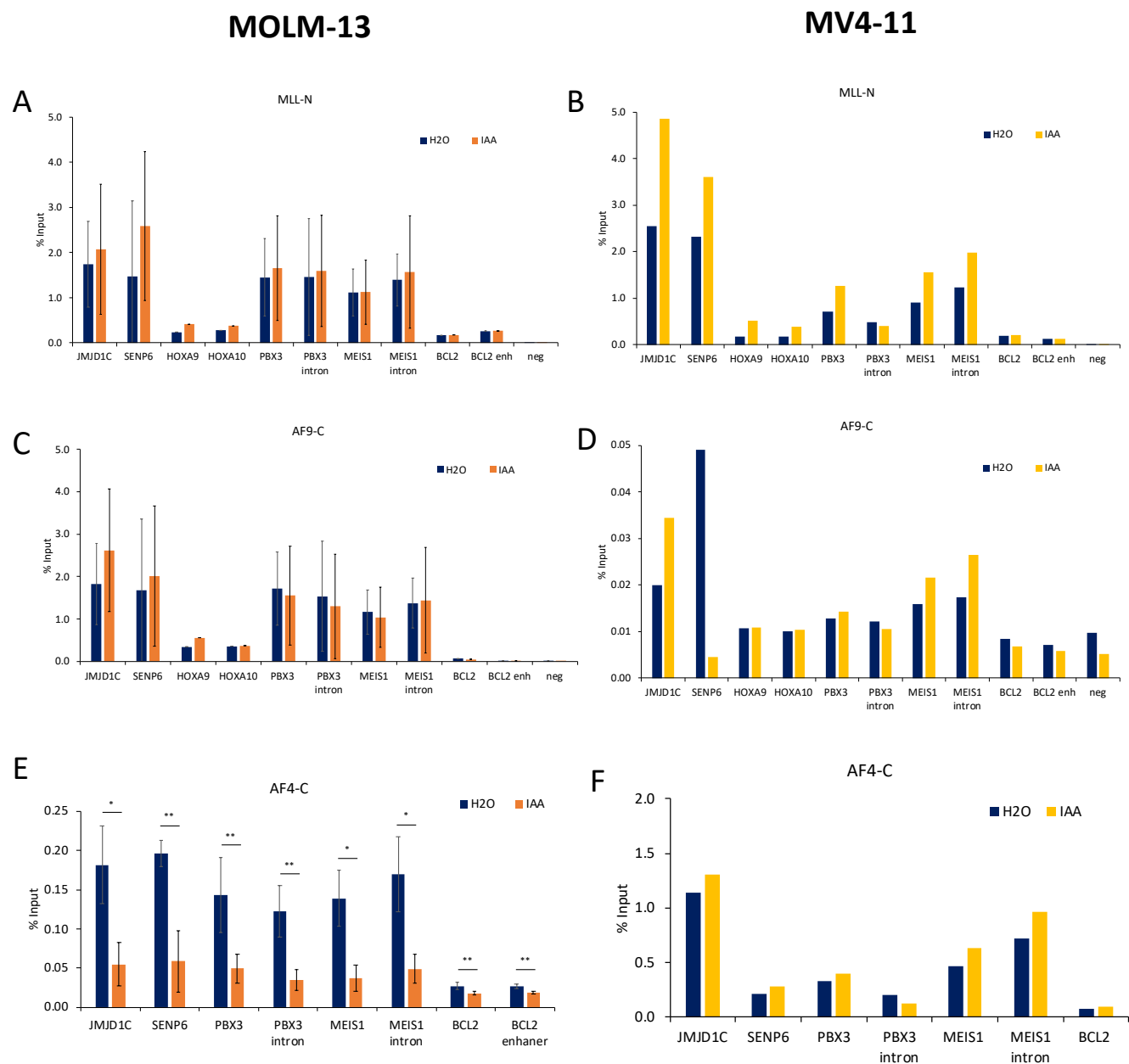


Figure 5.8 MLL, AF9 and AF4 occupancy at KAT7-bound loci following KAT7 loss

MOLM-13 and MV4-11 AID cells were treated with IAA for 24 hours. MLL-N terminal, AF9-C terminal and AF4-C terminal antibodies were used to probe for the binding of MLL-AF9 and MLL-AF4 fusion protein at various loci of selective MLL-fusion targets in MOLM-13 and MV4-11, respectively. Loci are promoter region unless indicated otherwise. A, C and E) ChIP-qPCR in MOLM-13, Mean \pm S.D., $n = 3$, two-tailed t -test (*, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$). B, D and F) ChIP-qPCR in MV4-11, $n = 1$.

5.9 Occupancy of BRD4 at MLL-fusion target loci is dependent on KAT7 in MOLM-13 but not in MV4-11

We subsequently investigated whether BRD4 occupancy at these target loci is affected by the loss of KAT7. This was motivated by several reasons. Firstly, BRD4 is a bromodomain protein, that is capable of “reading” histone acetylation marks, particularly lysine acetylation on H3 and H4 tails (Chiang, 2009; Devaiah et al., 2016). Secondly, from the findings illustrated in Chapter 4, KAT7 is responsible for the global acetylation of H3K14ac and H4K12ac, which are putative marks “read” by BRD4 (Chiang, 2009; Devaiah et al., 2016; Filippakopoulos and Knapp, 2012; Fiskus et al., 2014). Lastly, BRD4 interacts with and is a positive regulator of p-TEFb (a complex consisting of CDK9 and Cyclin T1/2), and p-TEFb is a member of various macromolecule complexes associated with MLL-fusion (Table 5.1) (Deshpande et al., 2012; Zeisig and So, 2015). Global BRD4 promoter occupancy is strongly correlated with the binding of KAT7 (Figure 5.9A). Similarly, at genetic loci co-bound by KAT7 and MLL-AF9 which are also downregulated in KAT7 knock-out show striking similarities in the binding profiles of these two proteins and BRD4 (Figure 5.9B).

24 hours of IAA treatment in MOLM-13 AID cells leads to a significant decrease in BRD4 occupancy at promoters of *JMJD1C*, *SENP6*, *PBX3*, *MEIS1* and *BCL2*, as well as introns of *PBX3* and *MEIS1* and the enhancer of *BCL2* (Figure 5.10A). With the exception of the *BCL2* promoter region, BRD4 binding is also slightly reduced at these loci in MV4-11 AID cells following 24 hours of IAA treatment (Figure 5.10C). We further assayed how RNA polymerase 2 serine 5 phosphorylation (RNA-PolII pS5) levels, a modification associated with transcription initiation and necessary for the transition to elongation (Ahn et al., 2004; Ferrari et al., 2014; Koch et al., 2008; Komarnitsky et al., 2000; Nojima et al., 2018; Soutourina, 2018), are affected upon loss of KAT7, and found striking decrease in all the target loci assessed— including promoter and introns in MOLM-13 (Figure 5.10B). However, in MV4-11, the changes in RNA-PolII pS5 levels at these loci are negligible (Figure 5.10D) and this may explain why these genes were not differentially expressed between KAT7 knock-out and control cells by RNA-seq experiments. Collectively, these findings suggest that BRD4 occupancy at these MLL-fusion target loci are dependent on KAT7 in MOLM-13 but not in MV4-11, potentially reflecting a different mechanism between MLL-AF9 and MLL-AF4 pathogenesis.

Most of the classical MLL-fusion targets assayed showed significant gene expression downregulation in MOLM-13 KAT7 knock-out relative to the empty control, but these genes are not differentially expressed in MV4-11 (Table 5.2). The weak changes in the occupancy of BRD4 and RNA-PolII pS5 in MV4-11 are consistent with the transcriptomic profiles of MV4-11. In particular, among the loci

assayed, only *JMJD1C* showed a significantly reduced expression in knock-out compared to the wild-type counterpart. It is likely that *JMJD1C*, along with the other selected MLL-fusion target genes are not the effectors responsible for maintaining the proliferation of MV4-11. The expression of the MLL-AF4 spreading targets that are significantly downregulated upon KAT7 knock-out (*NUDT5*, *PTEN*, *FLT3*, *NUSAP1*, *RAB27A*, *MRPL33*, *ITGA4* and *SATB1*), which were not assayed in this study, may depend on KAT7-mediated BRD4 binding for its expression and warrants further investigation. Alternatively, KAT7 in MV4-11 may have roles that are unrelated to BRD4 recruitment, such as the recruitment of other proteins via a different mechanism. Furthermore, KAT7 may be regulating transcription elongation rather than transcription initiation in MV4-11 and assaying RNA-PolII pS2 can elucidate this possibility.

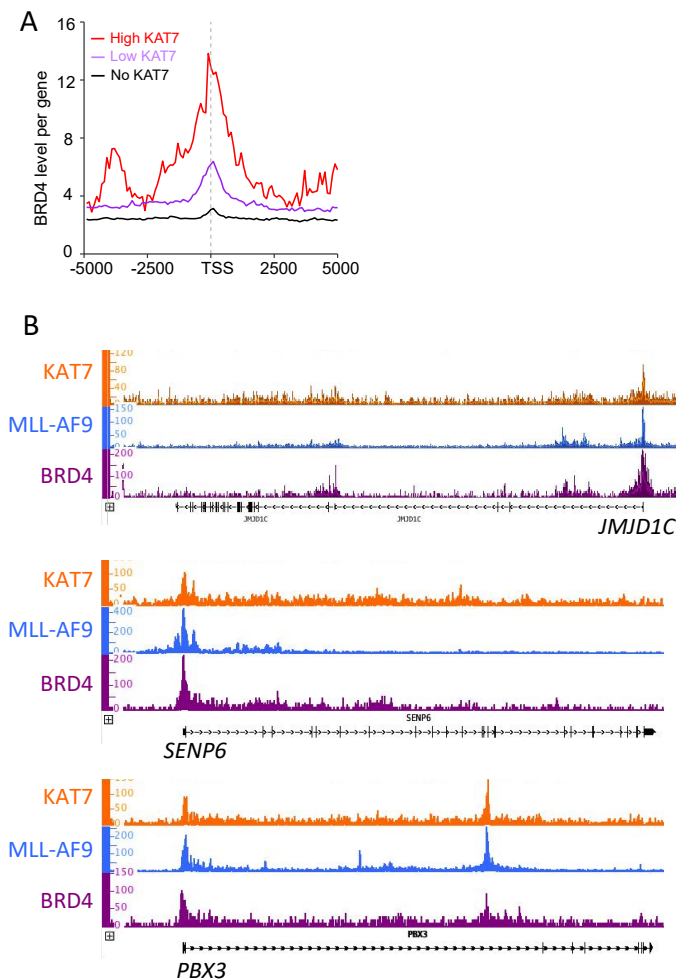


Figure 5.9 Co-localization of MLL-AF9, KAT7 and BRD4 at the promoter

Previously published BRD4 ChIP-seq in MV4-11 cell line (Gilan et al., 2016) and MLL-AF9 ChIP-seq in THP-1 (Kerry et al., 2017). A) Genome-wide correlation of KAT7 and BRD4 binding at the promoter.

B) Co-localization of KAT7, MLL-AF9 and BRD4 at the promoter of MLL-AF9 target genes JMJD1C, SENP6 and PBX3. Notably, the co-localization of these three proteins can also be seen at an intron of PBX3.

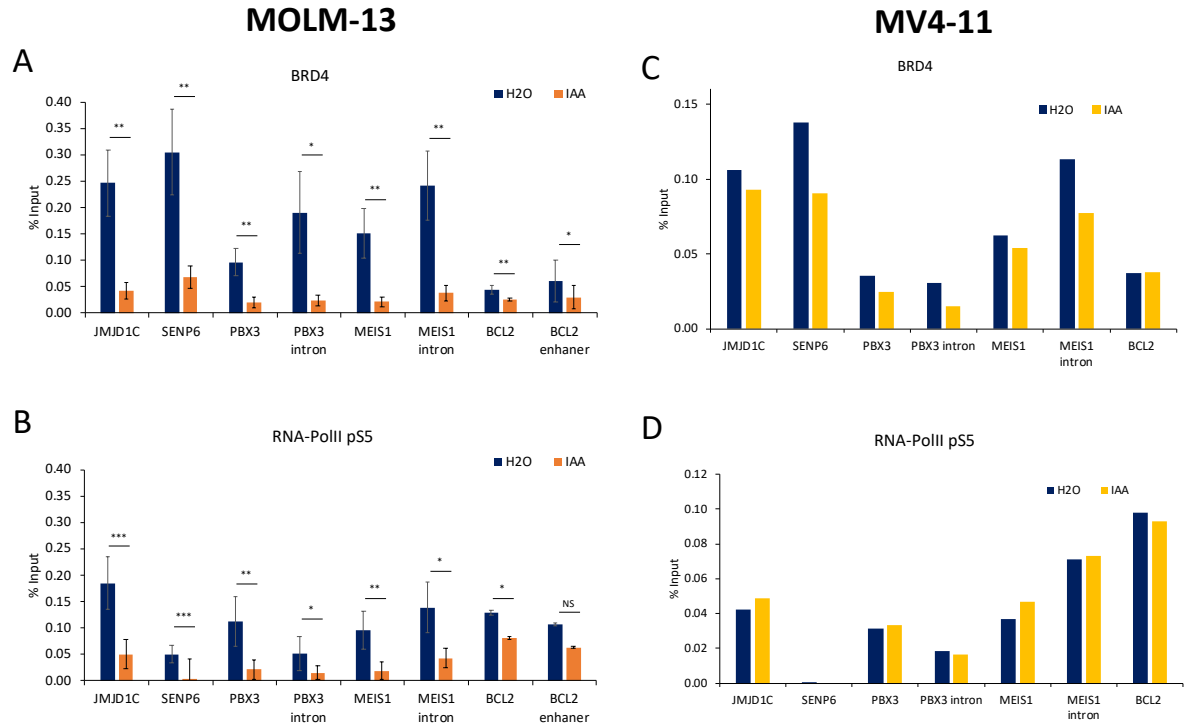


Figure 5.10 Occupancy of BRD4 and RNA Polymerase II Serine 5 phosphorylation at KAT7-bound loci following IAA treatment

MOLM-13 and MV4-11 AID cells were treated with IAA for 24 hours. The binding of BRD4 and RNA polymerase 2 serine 5 phosphorylation (RNA-PolII pS5) at the loci of various MLL-fusion targets in IAA treated compared to H₂O treated controls. Loci are promoter region unless indicated otherwise. A) BRD4 and B) RNA-PolII pS5 ChIP-qPCR in MOLM-13, Mean \pm S.D., $n = 3$, two-tailed t -test (NS, $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$). C) BRD4 and D) RNA-PolII pS5 ChIP-qPCR in MV4-11, $n = 1$.

5.10 Loci-specific changes in KAT7-mediated histone lysine acetylation levels upon IAA treatment

As KAT7 loss leads to a global decrease in histone acetylation levels such as H3K14ac, we wanted to investigate the local changes of various H3 and H4 lysine residues that are known targets of KAT7 at loci co-bound by KAT7 and spreading MLL-fusion.

24 hours after the addition of IAA, H3K14ac levels are significantly reduced at the promoter of *JMJD1C*, *SENP6* and *MEIS1* in MOLM-13 AID cells, (Figure 5.11A). Additionally, an intron of *PBX3* also showed a significant decrease in H3K14ac following the loss of KAT7 in MOLM-13 (Figure 5.11A). Data of one experimental replicate in MV4-11 suggests that H3K14ac levels are also reduced at the promoters of *JMJD1C*, *PBX3*, *MEIS1* and *BCL2*, as well as at introns of *PBX3* and *MEIS1* (Figure 5.11D). A similar trend is seen at these loci for H4K5ac in MV4-11 (Figure 5.11E). In MOLM-13, only the promoter of *JMJD1C* and the intron of *PBX3* showed a significant decrease in H4K5ac after KAT7 loss (Figure 5.11B). H4K8ac show a pattern of decrease at these loci in IAA treated MOLM-13 AID compared to control, but these changes are not statistically significant (Figure 5.11C). H4K8ac showed no changes at the promoter region of *JMJD1C* and *BCL2* in MV4-11 while an increase in this acetylation mark is observed at the intron of *MEIS1* after IAA treatment (Figure 5.11F). Promoter regions of *SENP6*, *PBX3* and *MEIS1*, as well as the intron for *PBX3*, are relatively decreased in IAA treated MV4-11 (Figure 5.11F). For MOLM-13, local changes in H3K14ac are consistent with the global depletion of this histone marks by immunoblot analysis.

Although not all statistically significant, KAT7 modified acetylation marks H4K5ac and H4K8ac are generally reduced at most loci assayed after the loss of KAT7 in both MOLM-13 and MV4-11. More repeats are required for MV4-11 ChIP-qPCR experiments to determine statistical significance. Notably, H4K5ac and H4K8ac signals are stronger compared to H3K14ac in both cell lines, which may be attributed to the quality of the H3K14ac antibody in recognizing the target or that the global levels of H4K5ac and H4K8ac are naturally more abundant than H3K14ac. It would be interesting to assay other KAT7-mediated histone lysine acetylation targets H3K23ac and H4K12ac to determine how all these histone marks change in response to KAT7 degradation in terms of genome-wide and loci-specific levels.

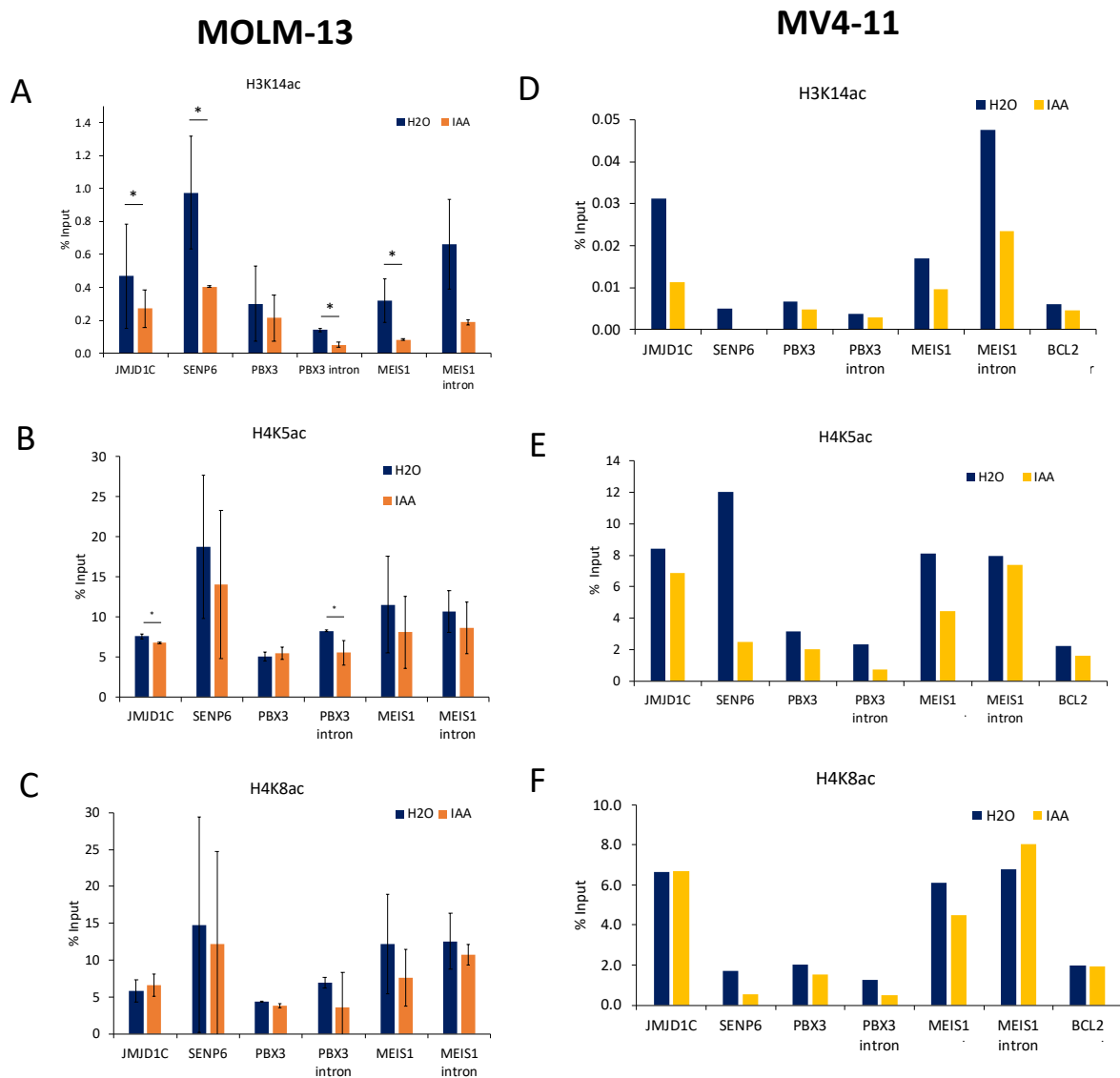


Figure 5.11 Changes in H3K14ac, H4K5ac and H4K8ac levels at KAT7-bound loci following depletion of KAT7 protein

MOLM-13 and MV4-11 AID cells were treated with IAA for 24 hours. Loci-specific changes in KAT7-mediated histone acetylation targets following the depletion of KAT7 protein. Loci are promoter region unless indicated otherwise. A) H3K14ac, B) H4K5ac and C) H4K8ac ChIP-qPCR in MOLM-13. Mean \pm S.D., $n = 3$, two-tailed t -test (NS, $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$). D) H3K14ac, E) H4K5ac and F) H4K8ac ChIP-qPCR in MV4-11, $n = 1$.

5.11 Working model for KAT7 mechanism of action in regulating the expression of MLL-AF9 target genes

KAT7 recruits BRD4 to the promoter of MLL-AF9 targets, presumably through the acetylation of lysine residues on histone H3 and H4 tails. This is associated with the promoter occupancy of AF4 protein. Binding of BRD4 and other MLL-fusion associated proteins such as AF4 drives the transcription of target genes (Figure 5.12A). In the absence of KAT7, acetylation on H3 and H4 tails are significantly reduced thereby result in the displacement of MLL-fusion associated proteins such as BRD4 and AF4 and consequently reduced expression of target genes (Figure 5.12B).

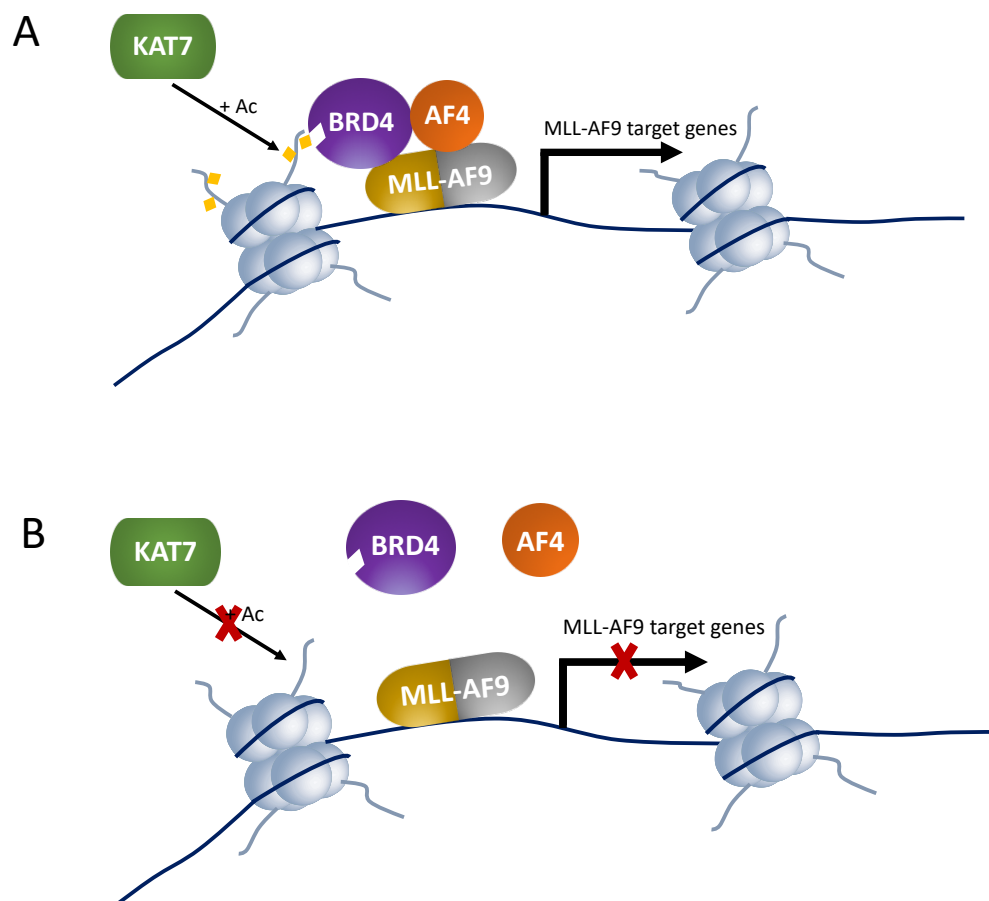


Figure 5.12 Working model for KAT7 mechanism of action in regulating the expression of MLL-AF9 target genes

5.12 MOLM-13 E508Q sensitizes cells to IBET-151 treatment

The data thus far provides strong evidence that KAT7 affects the binding of BRD4 at the promoter regions, as well as selective introns and enhancer, of many MLL-AF9 target genes. Therapeutically, it is interesting to investigate if the loss of KAT7 catalytic activity could sensitize AML cells to treatment with BET inhibitors such as IBET-151. We utilized the E508Q catalytic dead MOLM-13 cell line (with the presence of endogenous KAT7), in which the mutant would mimic the effects of potential KAT7 inhibitor. The viability of MOLM-13 E508Q mutant is significantly reduced, compared to GFP control, following 4 days of IBET-151 treatment (Figure 5.13). This is continued observed at 5 days of IBET-151 treatment. This enhanced sensitive in E508Q MOLM-13 cell lines compared to the control MOLM-13 cell suggests that BRD4 inhibitor-induced effect on viability is dependent on KAT7 activity, further strengthening the proposed model that KAT7 acts upstream of BRD4. This finding may encourage the combination treatment of BRD4 inhibitor with novel KAT7 inhibitor for MLL-fusion positive AML.

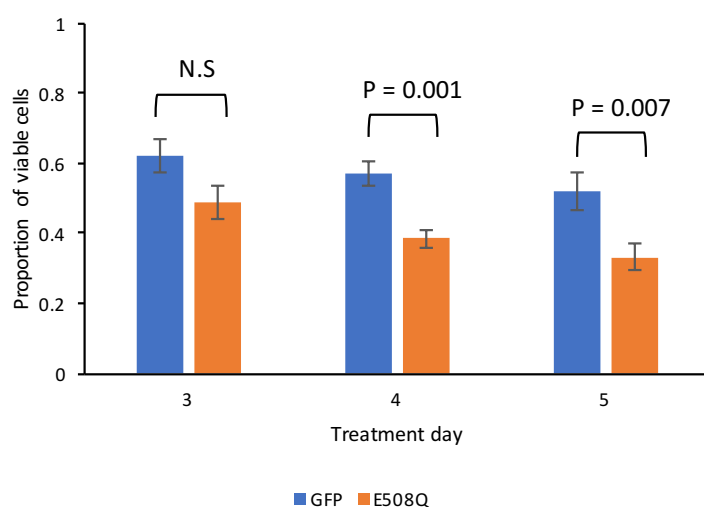


Figure 5.13 E508Q KAT7 mutant sensitizes MOLM-13 to IBET-151 treatment.

MOLM-13 E508Q mutant cells and MOLM-13 GFP cells were treated with 300 nM of IBET-151 for 3, 4 and 5 days and the viability were assayed at each time-point. Mean \pm S.D., $n = 3$, two-tailed t -test and p -value is indicated.

5.13 Discussion

Together, our finding suggests that KAT7 functions downstream of the MLL-fusion proteins where it is involved in the recruitment of MLL-AF9 associated accessory proteins such as BRD4 and AF4 to multiple target loci, presumably via the regulation of histone acetylation such as H3K14ac. The AID system mimics a KAT7 degrader as oppose to an enzymatic inhibitor of KAT7, therefore, the potential physical association of KAT7 at the promoter to facilitate the binding or recruitment of BRD4 cannot be eliminated. However, we have shown in Chapter 4 that the catalytic activity of KAT7 is essential for maintaining leukemogenesis in MOLM-13, which strongly implies that the catalytic function of KAT7 is responsible. Currently, it is not possible to distinguish whether histone or non-histone acetylation is functionally important. This is partly due to the lack of study on potential non-histone target of KAT7. Therefore, it would be valuable to perform a comprehensive analysis of the acetylated proteome (acetylome) before and after the loss of KAT7 protein to examine the possible protein lysine acetylation functions of KAT7. Furthermore, it would be interesting to investigate whether KAT7 forms protein complexes with MLL-AF9 and/or BRD4 at the promoter of target genes. Intriguingly, KAT7 was found to be co-purified with seven distinct MLL-fusion— MLL-AF9, MLL-AF4, MLL-ENL, MLL-AF1p, MLL-CBP, MLL-EEN and MLL-GAS7 (Skucha et al., 2018).

Our study strengthens previous observations that histone acetylation at the promoter of transcriptionally active genes facilitates the binding of BRD4 and the subsequent phosphorylation of RNA PolII (Dawson et al., 2011; Roe et al., 2015; Sheikh and Akhtar, 2018; Zhang et al., 2012). Data on the occupancy of total RNA PolII and RNA PolII serine 2 phosphorylation following KAT7 loss will elucidate further if KAT7 is involved in pre-initiation complex (PIC) binding and transcription elongation respectively. Importantly, our findings suggest that the reduction in the promoter binding of BRD4 following KAT7 loss is not associated with the change in occupancy of MLL-AF9 at the promoter and that the binding of MLL-AF9 without BRD4 is insufficient for maintaining the expression of its target genes, especially those highly bound by KAT7.

Interestingly, an intron region of *PBX3* is not only co-bound by KAT7, BRD4 and MLL-AF9, but the BRD4 binding at this locus is also significantly reduced in MOLM-13 following the depletion of KAT7 protein. Also, BRD4 is known to occupy enhancer regions (Dawson et al., 2014; Lovén et al., 2013) and whether KAT7 plays a role in this aspect can be investigated. The extent of KAT7 in regulating transcription via other DNA elements such as introns and enhancers warrants further investigation, in particular, the functional role of HAT occupancy, and the corresponding histone acetylation marks at introns is not well-characterized. Studies have suggested HATs and HDACS are co-enriched at

promoter and introns of transcriptionally active genes (Wang et al., 2009b; Yan Matthew S. and Marsden Philip A., 2015). H3K14ac, among other histone acetylation marks, have the majority of peaks at intron and distal intergenic regions, more than those found at gene promoters (Karmodiya et al., 2012; Yan Matthew S. and Marsden Philip A., 2015). KAT7 may additionally control gene expression via the binding at introns. In support of this notion, a recent study on endothelial cell (EC) gene regulation reported that the expression of VEGFR-2 and other EC-related genes is dependent on KAT7-mediated H3K14ac/pan-H4 acetylation in the intragenic region (Yan et al., 2018).

Despite the binding of KAT7 at a wide range of loci across the genome, only a subset of genes is downregulated upon KAT7 knock-out. This suggests that histone acetylation-mediated transcription regulation can be compensated for by other HATs at the majority of KAT7-bound loci. The dynamic occupancy of specific HATs and HDACs at a given genomic region likely governs the extent of BRD4 recruitment. BRD4 is a global regulator of transcription and is essential in a wide range of cells (Roe and Vakoc, 2016), however, KAT7 does not appear to be broadly essential at least among AML cell lines, suggesting that the mechanisms by which BRD4 is recruited to target loci can be context-specific. In the context of MLL-AF9 driven transcription, BRD4 occupancy at the promoter is strongly dependent on KAT7.

Chapter 6: Discussion and future work

Despite advances in understanding its genomics and molecular pathogenesis, AML remains fatal for the majority of the patients (Ferrara and Schiffer, 2013) and mainstream therapies have not changed significantly for several decades (Evans et al., 1961). Amongst different AML subtypes, cases driven by the MLL-fusion genes continue to represent a poor prognosis category (Krivtsov and Armstrong, 2007; Papaemmanuil et al., 2016; Saultz and Garzon, 2016) and while there are recent developments in the field (Albrecht et al., 2016; Berthon et al., 2016; Daigle et al., 2011; Dawson et al., 2011; Gallipoli et al., 2015; Grembecka et al., 2012; Klaus et al., 2014; Zuber et al., 2011) clinical progress is still lacking, emphasising the need for new therapies.

Here, we demonstrate that KAT7 represent a putative novel therapeutic target for MLL-fusion AML and give insights into its function in the maintenance of these leukemias, mainly in MLL-AF9 positive AML model. KAT7 is required for the recruitment of BRD4 and accessory proteins associated with transcription elongation complex such as AF4 to the promoter of MLL-AF9 target genes. Following reports of KAT7 as an AML-specific vulnerability from genome-wide CRISPR-Cas9 screening studies (Tzelepis et al., 2016; Wang et al., 2017), we are the first to decipher the underlying molecular mechanism of essentiality. Although no mutations in *KAT7* have been found in AML cases, our findings accentuate KAT7 as a vulnerability that could be therapeutically exploited in AML.

We are the first to elucidate a mechanistic link between KAT7 and BRD4. BRD4 is thought to be recruited to the promoter and enhancer regions genome-wide by lineage-specific transcription factors (Roe and Vakoc, 2016; Roe et al., 2015). Through the catalytic function of HATs such as p300 and CBP, HATs facilitate the recruitment of BRD4 by acetylation of both histone proteins and transcription factors (Roe and Vakoc, 2016; Roe et al., 2015). Interestingly, a study by Roe et al. investigated the essentiality of 17 mammalian HATs, KAT7 among one of them, in RN2 cells (murine AML with MLL-AF9 and NRAS^{G12D} mutations). Although *KAT7* was essential for the proliferation of RN2 cells, its transcriptomics profile following shRNA-mediated knock-down is inversely correlated with the transcriptomics of *BRD4* knock-down cells, suggesting that KAT7 perform regulatory function in opposition to BRD4 (Roe et al., 2015). Instead, the gene expression of *CBP* and *p300* knock-down resembles that of *BRD4* knock-down. Inhibition of CBP/p300 catalytic activity or knock-down of *p300* resulted in a similar reduction of BRD4 binding at the genome-wide promoter and enhancer regions. Although we did not compare the global changes in gene expression upon KAT7 loss and BRD4 loss in our human AML models, given that BRD4 promoter occupancy at MLL-AF9 target loci is dependent on KAT7, ablation of KAT7 and BRD4 would predictively lead to similar mRNA changes. Conversely,

Roe et al., did not investigate whether loss of KAT7 resulted in reduced BRD4 occupancy in RN2 cell lines. The potential discrepancy of findings between our studies and the study by Roe et al. may be attributed to the difference between human and mice hematopoietic system. The presence of other oncogenic mutations in each of the models may also contribute to the identification of different effector HATs.

Pharmacological inhibition of BET proteins, including BRD4, is an effective pre-clinical treatment for MLL-fusion leukemia (Chaidos et al., 2015; Dawson et al., 2011). Our studies provide evidence that loss of KAT7 protein results in the reduction of BRD4 occupancy at the promoter region of many canonical MLL-AF9 target loci. When using the AID protein-degradation system to study the molecular mechanism, we cannot eliminate the possibility that physical interaction of KAT7 is involved in the recruitment of BRD4 to these target loci. Given that potent and specific small molecules inhibiting the catalytic activity of KAT7 is not currently available, it is challenging to directly address this possibility due to the scalability of the existing experiment set up. However, we have shown through catalytic dead E508Q mutant KAT7 that the acetyltransferase activity of KAT7 is essential for leukemic maintenance in the MLL-AF9 context. It would be interesting to see if KAT7 physically associates with BRD4 and/or MLL-AF9 by co-immunoprecipitation experiments. Notably, our experiments conducted in this study does not distinguish between histone acetylation or protein acetylation. It would be interesting to ascertain if KAT7 also acetylates hematopoietic transcription factors, as with function identified for CBP/p300 (Roe et al., 2015). We cannot firmly conclude that BRD4 occupancy is mediated by KAT7-dependent acetylation, but the evidence collectively strongly suggests this is the case.

Yet to be addressed is how KAT7 genome-wide binding is regulated in the context of AML, particularly whether KAT7 is recruited by hematopoietic transcription factors important in MLL-fusion leukemogenesis. Although we did not find MLL-AF9 or MLL-AF4 occupancy to be significantly affected at various shared target loci after KAT7 protein degradation, we did observe a marginal increase in MLL-N and AF9-C and AF4-C occupancy. Furthermore, top KAT7 bound promoters are also spreading targets of MLL-AF9/MLL-AF4, suggesting that MLL-fusion may be selectively regulating KAT7 occupancy at these shared target loci. If MLL-fusion protein is involved in the recruitment of KAT7 to its target genes, a hypothetical theory could be the existence of a compensatory mechanism in which MLL-fusion chromatin binding increases in response to KAT7 loss in an attempt to “restore” the default levels of KAT7 occupancy at the target loci. A possible way to investigate whether MLL-fusion regulates KAT7 binding is to inhibit the chromatin binding of MLL-fusion via small molecules that

target the MENIN-MLL-fusion interactions or knock-down of MLL-AF9 by siRNA (Grembecka et al., 2012; Shi et al., 2012) followed by KAT7 ChIP-Seq or ChIP-qPCR.

It is possible that the chromatin binding of KAT7 and MLL-fusion is independent of one another, but the colocalization may depend on the same factors such as the presence of H3K36me3 and/or H3K4me3. As described in Chapter 1, depending on the proteins in which KAT7 associates with, the resulting protein complex may preferentially localize at H3K36me3 or H3K4me3. H3K36me3 can be recognized by the PWWP domain of LEDGF, a protein important for MLL-fusion target recognition (Eidahl et al., 2013; Okuda et al., 2014; Yokoyama, 2015, 2017; Zhu et al., 2016), whilst H3K4me3 is a mark “written” by wild-type MLL protein and the wild-type MLL is essential for MLL-AF9 pathogenesis (Krivtsov and Armstrong, 2007; Milne et al., 2010; Rao and Dou, 2015; Thiel et al., 2010; Wang et al., 2009a).

Various phenotypes reflected differences in response to KAT7 loss between MLL-AF9 and MLL-AF4 models, notably in terms of the ability to induce differentiation and the changes in occupancy of BRD4. MLL-fusion can association with distinct biochemical complexes that comprise of different protein components (Ballabio and Milne, 2012, 2014; Deshpande et al., 2012; Zeisig and So, 2015), as mentioned in Chapter 5. Our findings suggest that KAT7 is not involved in BRD4 recruitment in MV4-11 driven by the MLL-AF4 oncogene. Furthermore, KAT7 is an essential gene in OCI-AML2 (possessing MLL-AF6) and the genetic ablation of KAT7 induces differentiation and apoptosis in this cell line. The molecular mechanism by which KAT7 regulates the gene expression of MLL-AF6 transform cells remains elusive. In contrast to MLL-AF4 and MLL-AF9 oncoprotein, MLL-AF6 fusion protein homodimerizes and recruits DOT1L complex through an alternative mechanism that does not involve interaction with any transcriptional component (Ballabio and Milne, 2012, 2014; Deshpande et al., 2013; Liedtke et al., 2010; Manara et al., 2014; Yokoyama, 2015; Yokoyama et al., 2010). Regardless of the mechanisms of transformation by the respective MLL-fusion proteins, KAT7 is a shared vulnerability among our AML models with this oncogene. Therefore, the mechanisms of action of KAT7 in MLL-fusion other than MLL-AF9 remains an interesting and important avenue for further studies. It is also important to replicate the mechanistic study in another MLL-AF9 model such as THP-1 or Nomo-1 to ensure that this is a general mechanism common to MLL-AF9 leukemogenesis.

KAT7 does not appear to be involved in the haematopoiesis of human primary CD34+ progenitor and stem cells (HPSCs) or T-cells (Ting et al., 2018), which suggests that targeting *KAT7* may not have detrimental effects on the development of normal blood cells. Critically lacking in this study are data on the effects of KAT7 loss on primary human AML cells, particularly of the MLL-fusion subtype. The function of KAT7 in other cell types also needs to be addressed comprehensively through appropriate

in vivo models. The presence of KAT7 inhibitor and/or KAT7 degrader will facilitate the efforts in addressing these questions. Encouragingly, the recent development of inhibitors targeting KAT6A/B has shown promising tumour growth inhibitory effects *in vitro* and *in vivo* (Baell et al., 2018), confirming direct targeting of MYST family of HATs as a valid therapeutic approach.

Collectively, our findings revealed that KAT7 acts upstream of BRD4 to recruit MLL-fusion associated machineries to the promoter of MLL-AF9 target genes and may be an alternative therapeutic strategy for MLL-fusion leukemia. We anticipate that our work will motivate the development of potent small-molecules that are highly specific against KAT7 and highlight KAT7 as a potential novel therapeutic target for MLL-fusion AML.

References

- Adachi, N., Kimura, A., and Horikoshi, M. (2002). A Conserved Motif Common to the Histone Acetyltransferase Esa1 and the Histone Deacetylase Rpd3. *J. Biol. Chem.* 277, 35688–35695.
- Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of Serine 2 within the RNA Polymerase II C-Terminal Domain Couples Transcription and 3' End Processing. *Molecular Cell* 13, 67–76.
- Albrecht, B.K., Gehling, V.S., Hewitt, M.C., Vaswani, R.G., Côté, A., Leblanc, Y., Nasveschuk, C.G., Bellon, S., Bergeron, L., Campbell, R., et al. (2016). Identification of a Benzoisoxazoloazepine Inhibitor (CPI-0610) of the Bromodomain and Extra-Terminal (BET) Family as a Candidate for Human Clinical Trials. *J. Med. Chem.* 59, 1330–1339.
- Alinari, L., Mahasen, K. V., Yan, F., Karkhanis, V., Chung, J.-H., Smith, E.M., Quinion, C., Smith, P.L., Kim, L., Patton, J.T., et al. (2015). Selective inhibition of protein arginine methyltransferase 5 blocks initiation and maintenance of B-cell transformation. *Blood* 125, 2530–2543.
- Ashkar, S.E., Schwaller, J., Pieters, T., Goossens, S., Demeulemeester, J., Christ, F., Belle, S.V., Juge, S., Boeckx, N., Engelman, A., et al. (2018). LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. *Blood* 131, 95–107.
- Aubrey, B.J., Kelly, G.L., Janic, A., Herold, M.J., and Strasser, A. (2018). How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death and Differentiation* 25, 104–113.
- Avvakumov, N., and Côté, J. (2007). The MYST family of histone acetyltransferases and their intimate links to cancer. *Oncogene* 26, 5395–5407.
- Avvakumov, N., Lalonde, M.-E., Saksouk, N., Paquet, E., Glass, K.C., Landry, A.-J., Doyon, Y., Cayrou, C., Robitaille, G.A., Richard, D.E., et al. (2012). Conserved Molecular Interactions within the HBO1 Acetyltransferase Complexes Regulate Cell Proliferation. *Mol. Cell. Biol.* 32, 689–703.
- Ayton, P.M., and Cleary, M.L. (2001). Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene* 20, 5695–5707.
- Baell, J.B., Leaver, D.J., Hermans, S.J., Kelly, G.L., Brennan, M.S., Downer, N.L., Nguyen, N., Wichmann, J., McRae, H.M., Yang, Y., et al. (2018). Inhibitors of histone acetyltransferases KAT6A/B induce senescence and arrest tumour growth. *Nature* 560, 253.
- Bagger, F.O., Kinalis, S., and Rapin, N. (2019). BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Res* 47, D881–D885.
- Ballabio, E., and Milne, T.A. (2012). Molecular and Epigenetic Mechanisms of MLL in Human Leukemogenesis. *Cancers (Basel)* 4, 904–944.
- Ballabio, E., and Milne, T.A. (2014). Epigenetic control of gene expression in leukemogenesis: Cooperation between wild type MLL and MLL fusion proteins. *Molecular & Cellular Oncology* 1, e955330.
- Barth, T.K., and Imhof, A. (2010). Fast signals and slow marks: the dynamics of histone modifications. *Trends in Biochemical Sciences* 35, 618–626.

- Bassett, A.R., Kong, L., and Liu, J.L. (2015). A genome-wide CRISPR library for high-throughput genetic screening in drosophila cells. *Journal of Genetics and Genomics* 42, 301–309.
- Berndsen, C.E., Albaugh, B.N., Tan, S., and Denu, J.M. (2007). Catalytic mechanism of a MYST family histone acetyltransferase. *Biochemistry* 46, 623–629.
- Berthon, C., Raffoux, E., Thomas, X., Vey, N., Gomez-Roca, C., Yee, K., Taussig, D.C., Rezai, K., Roumier, C., Herait, P., et al. (2016). Bromodomain inhibitor OTX015 in patients with acute leukaemia: a dose-escalation, phase 1 study. *The Lancet Haematology* 3, e186–e195.
- Bhayat, F., Das-Gupta, E., Smith, C., McKeever, T., and Hubbard, R. (2009). The incidence of and mortality from leukaemias in the UK: a general population-based study. *BMC Cancer* 9, 252.
- Birke, M., Schreiner, S., García-Cuellar, M.-P., Mahr, K., Titgemeyer, F., and Slany, R.K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res* 30, 958–965.
- Borkin, D., He, S., Miao, H., Kempinska, K., Pollock, J., Chase, J., Purohit, T., Malik, B., Zhao, T., Wang, J., et al. (2015). Pharmacologic inhibition of the menin-MLL interaction blocks progression of MLL leukemia in vivo. *Cancer Cell* 27, 589–602.
- Borrow, J., Stanton, V.P., Andresen, J.M., Becher, R., Behm, F.G., Chaganti, R.S.K., Civin, C.I., Disteche, C., Dubé, I., Frischauf, A.M., et al. (1996). The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nature Genetics* 14, 33.
- Bose, D.A., Donahue, G., Reinberg, D., Shiekhatar, R., Bonasio, R., and Berger, S.L. (2017). RNA binding to CBP stimulates histone acetylation and transcription. *Cell* 168, 135-149.e22.
- Caenepeel, S., Brown, S.P., Belmontes, B., Moody, G., Keegan, K.S., Chui, D., Whittington, D.A., Huang, X., Poppe, L., Cheng, A.C., et al. (2018). AMG 176, a Selective MCL1 Inhibitor, Is Effective in Hematologic Cancer Models Alone and in Combination with Established Therapies. *Cancer Discov* 8, 1582–1597.
- Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how and why? *Mol Cell* 49.
- Carapeti, M., Aguiar, R.C.T., Goldman, J.M., and Cross, N.C.P. (1998). A Novel Fusion Between MOZ and the Nuclear Receptor Coactivator TIF2 in Acute Myeloid Leukemia. *Blood* 91, 3127–3133.
- Caslini, C., Yang, Z., El-Osta, M., Milne, T.A., Slany, R.K., and Hess, J.L. (2007). Interaction of MLL Amino Terminal Sequences with Menin Is Required for Transformation. *Cancer Res* 67, 7275–7283.
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov* 2, 401–404.
- Čermaková, K., Tesina, P., Demeulemeester, J., El Ashkar, S., Méreau, H., Schwaller, J., Řezáčová, P., Veverka, V., and De Rijck, J. (2014). Validation and structural characterization of the LEDGF/p75-MLL interface as a new target for the treatment of MLL-dependent leukemia. *Cancer Research* 74, 5139–5151.

- Chaidos, A., Caputo, V., and Karadimitris, A. (2015). Inhibition of bromodomain and extra-terminal proteins (BET) as a potential therapeutic approach in haematological malignancies: emerging preclinical and clinical evidence. *Ther Adv Hematol* 6, 128–141.
- Chan-Penebre, E., Kuplast, K.G., Majer, C.R., Boriack-Sjodin, P.A., Wigle, T.J., Johnston, L.D., Rioux, N., Munchhof, M.J., Jin, L., Jacques, S.L., et al. (2015). A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. *Nature Chemical Biology* 11, 432–437.
- Chen, C.-W., and Armstrong, S.A. (2015). Targeting DOT1L and HOX gene expression in MLL-rearranged leukemia and beyond. *Experimental Hematology* 43, 673–684.
- Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R., et al. (2015). Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* 160, 1246–1260.
- Chen, X., Liu, G., and Leffak, M. (2013). Activation of a human chromosomal replication origin by protein tethering. *Nucleic Acids Res* 41, 6460–6474.
- Chen, Z., Zhou, L., Wang, L., Kazobinka, G., Zhang, X., Han, X., Li, B., and Hou, T. (2017). HBO1 promotes cell proliferation in bladder cancer via activation of Wnt/ β -catenin signaling. *Molecular Carcinogenesis*.
- Chiang, C.-M. (2009). Brd4 engagement from chromatin targeting to transcriptional regulation: selective contact with acetylated histone H3 and H4. *F1000 Biol Rep* 1.
- Cierpicki, T., Risner, L.E., Grembecka, J., Lukasik, S.M., Popovic, R., Omonkowska, M., Shultis, D.D., Zeleznik-Le, N.J., and Bushweller, J.H. (2010). Structure of the MLL CXXC domain–DNA complex and its functional role in MLL-AF9 leukemia. *Nature Structural & Molecular Biology* 17, 62–68.
- Cotto, K.C., Wagner, A.H., Feng, Y.-Y., Kiwala, S., Coffman, A.C., Spies, G., Wollam, A., Spies, N.C., Griffith, O.L., and Griffith, M. (2018). DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. *Nucleic Acids Res* 46, D1068–D1073.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* 107, 21931–21936.
- Daigle, S.R., Olhava, E.J., Therkelsen, C.A., Majer, C.R., Sneeringer, C.J., Song, J., Johnston, L.D., Scott, M.P., Smith, J.J., Xiao, Y., et al. (2011). Selective Killing of Mixed Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. *Cancer Cell* 20, 53–65.
- Dawson, M.A., Prinjha, R.K., Dittman, A., Giotopoulos, G., Bantscheff, M., Chan, W.-I., Robson, S.C., Chung, C., Hopf, C., Savitski, M.M., et al. (2011). Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* 478, 529–533.
- Dawson, M.A., Gudgin, E.J., Horton, S.J., Giotopoulos, G., Meduri, E., Robson, S., Cannizzaro, E., Osaki, H., Wiese, M., Putwain, S., et al. (2014). Recurrent mutations, including NPM1c, activate a BRD4-dependent core transcriptional program in acute myeloid leukemia. *Leukemia* 28, 311–320.
- Decker, P.V., Yu, D.Y., Iizuka, M., Qiu, Q., and Smith, M.M. (2008). Catalytic-Site Mutations in the MYST Family Histone Acetyltransferase Esa1. *Genetics* 178, 1209–1220.
- Deng, M., Brägelmann, J., Kryukov, I., Saraiva-Agostinho, N., and Perner, S. (2017). FirebrowseR: an R client to the Broad Institute’s Firehose Pipeline. *Database (Oxford)* 2017.

- Deschler, B., and Lübbert, M. (2006). Acute myeloid leukemia: epidemiology and etiology. *Cancer* 107, 2099–2107.
- Deshpande, A.J., Bradner, J., and Armstrong, S.A. (2012). Chromatin modifications as therapeutic targets in MLL-rearranged leukemia. *Trends in Immunology* 33, 563–570.
- Deshpande, A.J., Chen, L., Fazio, M., Sinha, A.U., Bernt, K.M., Banka, D., Dias, S., Chang, J., Olhava, E.J., Daigle, S.R., et al. (2013). Leukemic transformation by the MLL-AF6 fusion oncogene requires the H3K79 methyltransferase Dot1l. *Blood* 121, 2533–2541.
- Devaiah, B.N., Case-Borden, C., Gegonne, A., Hsu, C.H., Chen, Q., Meerzaman, D., Dey, A., Ozato, K., and Singer, D.S. (2016). BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin. *Nature Structural and Molecular Biology* 23, nsmb.3228.
- DiNardo, C.D., Pratz, K., Pullarkat, V., Jonas, B.A., Arellano, M., Becker, P.S., Frankfurt, O., Konopleva, M., Wei, A.H., Kantarjian, H.M., et al. (2019). Venetoclax combined with decitabine or azacitidine in treatment-naïve, elderly patients with acute myeloid leukemia. *Blood* 133, 7–17.
- Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., et al. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481, 506–510.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Döhner, H., Weisdorf, D.J., and Bloomfield, C.D. (2015). Acute myeloid leukemia. *The New England Journal of Medicine* 373, 1136–1152.
- Döhner, H., Estey, E.H., Amadori, S., Appelbaum, F.R., Bu, T., Burnett, A.K., Fenaux, P., Grimwade, D., Larson, R.A., Lo-coco, F., et al. (2016). Diagnosis and management of acute myeloid leukemia in adults : recommendations from an international expert panel , on behalf of the European LeukemiaNet. *Blood* 115, 453–475.
- Dores, G.M., Devesa, S.S., Curtis, R.E., Linet, M.S., and Morton, L.M. (2012). Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. *Blood* 119, 34–43.
- Doyon, Y., Cayrou, C., Ullah, M., Landry, A.-J., Côté, V., Selleck, W., Lane, W.S., Tan, S., Yang, X.-J., and Côté, J. (2006). ING Tumor Suppressor Proteins Are Critical Regulators of Chromatin Acetylation Required for Genome Expression and Perpetuation. *Molecular Cell* 21, 51–64.
- Eidahl, J.O., Crowe, B.L., North, J.A., McKee, C.J., Shkriabai, N., Feng, L., Plumb, M., Graham, R.L., Gorelick, R.J., Hess, S., et al. (2013). Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Res* 41, 3924–3936.
- Erb, M.A., Scott, T.G., Li, B.E., Xie, H., Paulk, J., Seo, H.-S., Souza, A., Roberts, J.M., Dastjerdi, S., Buckley, D.L., et al. (2017). Transcription control by the ENL YEATS domain in acute leukaemia. *Nature* 543, 270–274.
- Esteyries, S., Perot, C., Adelaide, J., Imbert, M., Lagarde, A., Pautas, C., Olschwang, S., Birnbaum, D., Chaffanet, M., and Mozziconacci, M.-J. (2008). NCOA3, a new fusion partner for MOZ/MYST3 in M5 acute myeloid leukemia. *Leukemia* 22, 663–665.

- Evans, J.S., Musser, E.A., Mengel, G.D., Forsblad, K.R., and Hunter, J.H. (1961). Antitumor Activity of 1- β -D-Arabinofuranosylcytosine Hydrochloride. *Proceedings of the Society for Experimental Biology and Medicine* 106, 350–353.
- Feng, Y., Vlassis, A., Roques, C., Lalonde, M.-E., González-Aguilera, C., Lambert, J.-P., Lee, S.-B., Zhao, X., Alabert, C., Johansen, J.V., et al. (2016). BRPF3-HBO1 regulates replication origin activation and histone H3K14 acetylation. *The EMBO Journal* 35, 176–192.
- Ferrara, F., and Schiffer, C.A. (2013). Acute myeloid leukaemia in adults. *Lancet* 381, 484–495.
- Ferrari, F., Alekseyenko, A.A., Park, P.J., and Kuroda, M.I. (2014). Transcriptional control of a whole chromosome: emerging models for dosage compensation. *Nat Struct Mol Biol* 21, 118–125.
- Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., et al. (2010a). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* 18, 553–567.
- Figueroa, M.E., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P.J., Schifano, E., Booth, J., van Putten, W., Skrabanek, L., et al. (2010b). DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell* 17, 13–27.
- Filippakopoulos, P., and Knapp, S. (2012). The bromodomain interaction module. *FEBS Letters* 586, 2692–2704.
- Fiskus, W., Sharma, S., Qi, J., Valenta, J.A., Schaub, L.J., Shah, B., Peth, K., Portier, B.P., Rodriguez, M., Devaraj, S.G.T., et al. (2014). Highly active combination of BRD4 antagonist and histone deacetylase inhibitor against human acute myeloid leukemia (AML) cells. *Mol Cancer Ther* molcanther.0770.2013.
- Gallipoli, P., Giotopoulos, G., and Huntly, B.J.P. (2015). Epigenetic regulators as promising therapeutic targets in acute myeloid leukemia. *Ther Adv Hematol* 6, 103–119.
- Galmarini, C., Mackey, J., and Dumontet, C. (2001). Nucleoside analogues: mechanisms of drug resistance and reversal strategies. *Leukemia* 15, 875–890.
- Gang, E.J., Hsieh, Y.-T., Pham, J., Zhao, Y., Nguyen, C., Huantes, S., Park, E., Naing, K., Klemm, L., Swaminathan, S., et al. (2014). Small-molecule inhibition of CBP/catenin interactions eliminates drug-resistant clones in acute lymphoblastic leukemia. *Oncogene* 33, 2169–2178.
- Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013a). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* 6, pl1–pl1.
- Gao, X., Lin, J., Ning, Q., Gao, L., Yao, Y., Zhou, J., Li, Y., Wang, L., and Yu, L. (2013b). A histone acetyltransferase p300 inhibitor C646 induces cell cycle arrest and apoptosis selectively in AML1-ETO-positive AML cells. *PLoS ONE* 8, e55481.
- Garson, O.M., Hagemeijer, A., Sakurai, M., Reeves, B.R., Swansbury, G.J., Williams, G.J., Alimena, G., Arthur, D.C., Berger, R., and de la Chapelle, A. (1989). Cytogenetic studies of 103 patients with acute myelogenous leukemia in relapse. *Cancer Genetics and Cytogenetics* 40, 187–202.
- Georgiakaki, M., Chabbert-Buffet, N., Dasen, B., Meduri, G., Wenk, S., Rajhi, L., Amazit, L., Chauchereau, A., Burger, C.W., Blok, L.J., et al. (2006). Ligand-Controlled Interaction of Histone

Acetyltransferase Binding to ORC-1 (HBO1) with the N-Terminal Transactivating Domain of Progesterone Receptor Induces Steroid Receptor Coactivator 1-Dependent Coactivation of Transcription. *Mol Endocrinol* 20, 2122–2140.

Gilan, O., Lam, E.Y.N., Becher, I., Lugo, D., Cannizzaro, E., Joberty, G., Ward, A., Wiese, M., Fong, C.Y., Ftouni, S., et al. (2016). Functional interdependence of BRD4 and DOT1L in MLL leukemia. *Nature Structural and Molecular Biology* 23, nsmb.3249.

Girard, N., Bazille, C., Lhuissier, E., Benateau, H., Llombart-Bosch, A., Boumediene, K., and Bauge, C. (2014). 3-Deazaneplanocin A (DZNep), an inhibitor of the histone methyltransferase EZH2, induces apoptosis and reduces cell migration in chondrosarcoma cells. *PLoS ONE* 9, e98176.

Grembecka, J., He, S., Shi, A., Purohit, T., Muntean, A.G., Sorenson, R.J., Showalter, H.D., Murai, M.J., Belcher, A.M., Hartley, T., et al. (2012). Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. *Nature Chemical Biology* 8, 277–284.

Grimwade, D., Ivey, A., and Huntly, B.J.P. (2015). Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. *Blood* 127, 29–42.

Grossman, R.L., Heath, A.P., Ferretti, V., Varmus, H.E., Lowy, D.R., Kibbe, W.A., and Staudt, L.M. (2016). Toward a Shared Vision for Cancer Genomic Data. *New England Journal of Medicine* 375, 1109–1112.

Guenther, M.G., Lawton, L.N., Rozovskaia, T., Frampton, G.M., Levine, S.S., Volkert, T.L., Croce, C.M., Nakamura, T., Canaani, E., and Young, R.A. (2008). Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. *Genes Dev.* 22, 3403–3408.

Hart, T., Chandrashekar, M., Aregger, M., Steinhart, Z., Brown, K.R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., et al. (2015). High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* 163, 1515–1526.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone Modifications at Human Enhancers Reflect Global Cell Type-Specific Gene Expression. *Nature* 459, 108–112.

Hird, A.W., and Tron, A.E. (2019). Recent advances in the development of Mcl-1 inhibitors for cancer therapy. *Pharmacology & Therapeutics* 198, 59–67.

Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *PNAS* 109, E3350–E3357.

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science (New York, N.Y.)* 327, 167–170.

Ibrahimi, A., Velde, G.V., Reumers, V., Toelen, J., Thiry, I., Vandeputte, C., Vets, S., Deroose, C., Bormans, G., Baekelandt, V., et al. (2009). Highly Efficient Multicistronic Lentiviral Vectors with Peptide 2A Sequences. *Human Gene Therapy* 20, 845–860.

Iizuka, M., and Stillman, B. (1999). Histone Acetyltransferase HBO1 Interacts with the ORC1 Subunit of the Human Initiator Protein. *J. Biol. Chem.* 274, 23027–23034.

Iizuka, M., Matsui, T., Takisawa, H., and Smith, M.M. (2006). Regulation of Replication Licensing by Acetyltransferase Hbo1. *Molecular and Cellular Biology* 26, 1098–1108.

Iizuka, M., Sarmiento, O.F., Sekiya, T., Scrable, H., Allis, C.D., and Smith, M.M. (2008). Hbo1 Links p53-Dependent Stress Signaling to DNA Replication Licensing. *Molecular and Cellular Biology* 28, 140–153.

Iizuka, M., Takahashi, Y., Mizzen, C.A., Cook, R.G., Fujita, M., Allis, C.D., Frierson, H.F., Fukusato, T., and Smith, M.M. (2009). Histone acetyltransferase Hbo1: Catalytic activity, cellular abundance, and links to primary cancers. *Gene* 436, 108–114.

Iizuka, M., Susa, T., Takahashi, Y., Tamamori-Adachi, M., Kajitani, T., Okinaga, H., Fukusato, T., and Okazaki, T. (2013). Histone acetyltransferase Hbo1 destabilizes estrogen receptor α by ubiquitination and modulates proliferation of breast cancers. *Cancer Science* 104, 1647–1655.

IIZUKA, M., SUSU, T., TAMAMORI-ADACHI, M., OKINAGA, H., and OKAZAKI, T. (2017). Intrinsic ubiquitin E3 ligase activity of histone acetyltransferase Hbo1 for estrogen receptor α . *Proc Jpn Acad Ser B Phys Biol Sci* 93, 498–510.

Jin, Q., Yu, L.-R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.-E., Wang, C., Brindle, P.K., Dent, S.Y.R., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J* 30, 249–262.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* 1, 324–337.

Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* 13, 424.

Kerry, J., Godfrey, L., Repapi, E., Tapia, M., Blackledge, N.P., Ma, H., Ballabio, E., O’Byrne, S., Ponthan, F., Heidenreich, O., et al. (2017). MLL-AF4 Spreading Identifies Binding Sites that Are Distinct from Super-Enhancers and that Govern Sensitivity to DOT1L Inhibition in Leukemia. *Cell Reports* 18, 482–495.

Kim, J.H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K.Y., Kim, M.-K., Shin, B.A., and Choi, S.-Y. (2011). High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice. *PLoS One* 6.

King, A.D., Huang, K., Rubbi, L., Liu, S., Wang, C.-Y., Wang, Y., Pellegrini, M., and Fan, G. (2016). Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells. *Cell Rep* 17, 289–302.

Kitabayashi, I., Aikawa, Y., Yokoyama, A., Hosoda, F., Nagai, M., Kakazu, N., Abe, T., and Ohki, M. (2001). Fusion of MOZ and p300 histone acetyltransferases in acute monocytic leukemia with a t(8;22)(p11;q13) chromosome translocation. *Leukemia* 15, 89–94.

Klaus, C.R., Iwanowicz, D., Johnston, D., Campbell, C.A., Smith, J.J., Moyer, M.P., Copeland, R.A., Olhava, E.J., Scott, M.P., Pollock, R.M., et al. (2014). DOT1L Inhibitor EPZ-5676 Displays Synergistic Antiproliferative Activity in Combination with Standard of Care Drugs and Hypomethylating Agents in MLL-Rearranged Leukemia Cells. *J Pharmacol Exp Ther* 350, 646–656.

Knutson, S.K., Wigle, T.J., Warholic, N.M., Sneeringer, C.J., Allain, C.J., Klaus, C.R., Sacks, J.D., Raimondi, A., Majer, C.R., Song, J., et al. (2012). A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nature Chemical Biology* 8, 890–896.

- Koch, F., Jourquin, F., Ferrier, P., and Andrau, J.-C. (2008). Genome-wide RNA polymerase II: not genes only! *Trends in Biochemical Sciences* 33, 265–273.
- Kohlmann, A., Schoch, C., Dugas, M., Schnittger, S., Hiddemann, W., Kern, W., and Haferlach, T. (2005). New insights into *MLL* gene rearranged acute leukemias using gene expression profiling: shared pathways, lineage commitment, and partner genes. *Leukemia* 19, 953–964.
- Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M.D.C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nature Biotechnology* 32, 267–273.
- Komarnitsky, P., Cho, E.-J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* 14, 2452–2460.
- Kontro, M., Eldfors, S., Majumder, M.M., Parsons, A., Pemovska, T., Kallioniemi, O., Wennerberg, K., Heckman, C.A., and Porkka, K. (2015). BCL2-Inhibitors Target a Major Group of Newly-Diagnosed and Relapsed/Refractory Acute Myeloid Leukemia Ex Vivo. *Blood* 126, 2462–2462.
- De Kouchkovsky, I., and Abdul-Hay, M. (2016). ‘Acute myeloid leukemia: a comprehensive review and 2016 update.’ *Blood Cancer Journal* 6, e441.
- Krivtsov, A.V., and Armstrong, S.A. (2007). *MLL* translocations, histone modifications and leukaemia stem-cell development. *Nature Reviews Cancer* 7, 823–833.
- Krivtsov, A.V., Twomey, D., Feng, Z., Stubbs, M.C., Wang, Y., Faber, J., Levine, J.E., Wang, J., Hahn, W.C., Gilliland, D.G., et al. (2006). Transformation from committed progenitor to leukaemia stem cell initiated by *MLL*–AF9. *Nature* 442, 818–822.
- Kueh, A.J., Dixon, M.P., Voss, A.K., and Thomas, T. (2011). HBO1 Is Required for H3K14 Acetylation and Normal Transcriptional Activity during Embryonic Development. *Molecular and Cellular Biology* 31, 845–860.
- Kufe, D.W., Major, P.P., Egan, E.M., and Beardsley, G.P. (1980). Correlation of cytotoxicity with incorporation of ara-C into DNA. *The Journal of Biological Chemistry* 255, 8997–8900.
- Kühn, M.W.M., and Armstrong, S.A. (2015). Designed to Kill: Novel Menin-*MLL* Inhibitors Target *MLL*-Rearranged Leukemia. *Cancer Cell* 27, 431–433.
- Lalonde, M.-E., Avvakumov, N., Glass, K.C., Joncas, F.-H., Saksouk, N., Holliday, M., Paquet, E., Yan, K., Tong, Q., Klein, B.J., et al. (2013). Exchange of associated factors directs a switch in HBO1 acetyltransferase histone tail specificity. *Genes Dev.* 27, 2009–2024.
- Lavy, M., and Estelle, M. (2016). Mechanisms of auxin signaling. *Development* 143, 3226–3229.
- Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218.
- Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505.

- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Li, S., Garrett-Bakelman, F.E., Chung, S.S., Sanders, M.A., Hricik, T., Rapaport, F., Patel, J., Dillon, R., Vijay, P., Brown, A.L., et al. (2016a). Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. *Nature Medicine* 22, 792–799.
- Li, S., Mason, C.E., and Melnick, A. (2016b). Genetic and epigenetic heterogeneity in acute myeloid leukemia. *Current Opinion in Genetics & Development* 36, 100–106.
- Li, Z., He, S., and Look, A.T. (2019). The MCL1-specific inhibitor S63845 acts synergistically with venetoclax/ABT-199 to induce apoptosis in T-cell acute lymphoblastic leukemia cells. *Leukemia* 33, 262–266.
- Liang, G., Lin, J.C.Y., Wei, V., Yoo, C., Cheng, J.C., Nguyen, C.T., Weisenberger, D.J., Egger, G., Takai, D., Gonzales, F.A., et al. (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc Natl Acad Sci U S A* 101, 7357–7362.
- Liedtke, M., and Cleary, M.L. (2009). Therapeutic targeting of MLL. *Blood* 113, 6061–6068.
- Liedtke, M., Ayton, P.M., Somervaille, T.C.P., Smith, K.S., and Cleary, M.L. (2010). Self-association mediated by the Ras association 1 domain of AF6 activates the oncogenic potential of MLL-AF6. *Blood* 116, 63–70.
- Liesveld, J. (2012). Management of AML: Who do we really cure? *Leukemia Research* 36, 1475–1480.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15, 550.
- Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. *Cell* 153, 320–334.
- Lowenberg, B., Downing, J.R., and Burnett, A. (1999). Acute myeloid leukemia. *New England Journal of Medicine* 341, 1051–1062.
- Major, P.P., Egan, E.M., Beardsley, G.P., Minden, M.D., and Kufe, D.W. (1981). Lethality of human myeloblasts correlates with the incorporation of arabinofuranosylcytosine into DNA. *Proceedings of the National Academy of Sciences of the United States of America* 78, 3235–3239.
- Major, P.P., Egan, E.M., Herrick, D.J., and Kufe, D.W. (1982). Effect of ara-C incorporation on deoxyribonucleic acid synthesis in cells. *Biochemical Pharmacology* 31, 2937–2940.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology* 9, 467–477.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* 13, 722–736.

- Manara, E., Baron, E., Tregnago, C., Aveic, S., Bisio, V., Bresolin, S., Masetti, R., Locatelli, F., Basso, G., and Pigazzi, M. (2014). MLL-AF6 fusion oncogene sequesters AF6 into the nucleus to trigger RAS activation in myeloid leukemia. *Blood* 124, 263–272.
- Marschalek, R. (2011). Mechanisms of leukemogenesis by MLL fusion proteins. *British Journal of Haematology* 152, 141–154.
- Matano, M., Date, S., Shimokawa, M., Takano, A., Fujii, M., Ohta, Y., Watanabe, T., Kanai, T., and Sato, T. (2015). Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nature Medicine* 21, 256–262.
- Meyer, C., Hofmann, J., Burmeister, T., Gröger, D., Park, T.S., Emerenciano, M., Oliveira, M.P. de, Renneville, A., Villarese, P., Macintyre, E., et al. (2013). The *MLL* recombinome of acute leukemias in 2013. *Leukemia* 27, leu2013135.
- Meyer, C., Burmeister, T., Gröger, D., Tsaar, G., Fechina, L., Renneville, A., Sutton, R., Venn, N.C., Emerenciano, M., Pombo-de-Oliveira, M.S., et al. (2018). The *MLL* recombinome of acute leukemias in 2017. *Leukemia* 32, 273–284.
- Milne, T.A., Kim, J., Wang, G.G., Stadler, S.C., Basrur, V., Whitcomb, S.J., Wang, Z., Ruthenburg, A.J., Elenitoba-Johnson, K.S.J., Roeder, R.G., et al. (2010). Multiple Interactions Recruit MLL1 and MLL1 Fusion Proteins to the HOXA9 Locus in Leukemogenesis. *Molecular Cell* 38, 853–863.
- Minotti, G. (2004). Anthracyclines: molecular advances and pharmacologic developments in antitumor Activity and Cardiotoxicity. *Pharmacological Reviews* 56, 185–229.
- Miotto, B., and Struhl, K. (2008). HBO1 histone acetylase is a coactivator of the replication licensing factor Cdt1. *Genes Dev.* 22, 2633–2638.
- Miotto, B., and Struhl, K. (2010). HBO1 Histone Acetylase Activity Is Essential for DNA Replication Licensing and Inhibited by Geminin. *Molecular Cell* 37, 57–66.
- Miotto, B., and Struhl, K. (2011). JNK1 Phosphorylation of Cdt1 Inhibits Recruitment of HBO1 Histone Acetylase and Blocks Replication Licensing in Response to Stress. *Molecular Cell* 44, 62–71.
- Mishima, Y., Miyagi, S., Saraya, A., Negishi, M., Endoh, M., Endo, T.A., Toyoda, T., Shinga, J., Katsumoto, T., Chiba, T., et al. (2011). The Hbo1-Brd1/Brpf2 complex is responsible for global acetylation of H3K14 and required for fetal liver erythropoiesis. *Blood* 118, 2443–2453.
- Montero, J., and Letai, A. (2018). Why do BCL-2 inhibitors work and where should we use them in the clinic? *Cell Death and Differentiation* 25, 56–64.
- Mrózek, K., Heerema, N.A., and Bloomfield, C.D. (2004). Cytogenetics in acute leukemia. *Blood Reviews* 18, 115–136.
- Muntean, A.G., and Hess, J.L. (2012). The Pathogenesis of Mixed-Lineage Leukemia. *Annual Review of Pathology: Mechanisms of Disease* 7, 283–301.
- Muntean, A.G., Tan, J., Sitwala, K., Huang, Y., Bronstein, J., Connelly, J.A., Basrur, V., Elenitoba-Johnson, K.S.J., and Hess, J.L. (2010). The PAF Complex Synergizes with MLL Fusion Proteins at HOX Loci to Promote Leukemogenesis. *Cancer Cell* 17, 609–621.
- Musselman, C.A., Lalonde, M.-E., Côté, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. *Nature Structural & Molecular Biology* 19, 1218–1227.

- Natsume, T., Kiyomitsu, T., Saga, Y., and Kanemaki, M.T. (2016). Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Reports* 15, 210–218.
- Neff, T., and Armstrong, S.A. (2013). Recent progress toward epigenetic therapies: the example of mixed lineage leukemia. *Blood* 121, 4847–4853.
- Nguyen, A.T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. *Genes Dev.* 25, 1345–1358.
- Nishimura, K., and Kanemaki, M.T. (2014). Rapid Depletion of Budding Yeast Proteins via the Fusion of an Auxin-Inducible Degron (AID). *Current Protocols in Cell Biology* 64, 20.9.1–20.9.16.
- Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* 6, 917–922.
- Nojima, T., Rebelo, K., Gomes, T., Grosso, A.R., Proudfoot, N.J., and Carmo-Fonseca, M. (2018). RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. *Molecular Cell* 72, 369–379.e4.
- Okuda, H., Kawaguchi, M., Kanai, A., Matsui, H., Kawamura, T., Inaba, T., Kitabayashi, I., and Yokoyama, A. (2014). MLL fusion proteins link transcriptional coactivators to previously active CpG-rich promoters. *Nucleic Acids Res* 42, 4241–4256.
- Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V.I., Paschka, P., Roberts, N.D., Potter, N.E., Heuser, M., Thol, F., Bolli, N., et al. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. *New England Journal of Medicine* 374, 2209–2221.
- Perini, G.F., Ribeiro, G.N., Pinto Neto, J.V., Campos, L.T., and Hamerschlag, N. (2018). BCL-2 as therapeutic target for hematological malignancies. *Journal of Hematology & Oncology* 11, 65.
- Pollyea, D.A., de Botton, S., Fathi, A.T., Stein, E.M., Tallman, M.S., Agresta, S., Bowden, C., Fan, B., Prah, M., Yang, H., et al. (2014). 1LBA Clinical safety and activity in a phase I trial of AG-120, a first in class, selective, potent inhibitor of the IDH1-mutant protein, in patients with IDH1 mutant positive advanced hematologic malignancies. *European Journal of Cancer* 50, 195.
- Pollyea, D.A., Stevens, B.M., Jones, C.L., Winters, A., Pei, S., Minhajuddin, M., D'Alessandro, A., Culp-Hill, R., Riemondy, K.A., Gillen, A.E., et al. (2018). Venetoclax with azacitidine disrupts energy metabolism and targets leukemia stem cells in patients with acute myeloid leukemia. *Nat Med* 24, 1859–1866.
- Prange, K.H.M., Mandoli, A., Kuznetsova, T., Wang, S.-Y., Sotoca, A.M., Marneth, A.E., van der Reijden, B.A., Stunnenberg, H.G., and Martens, J.H.A. (2017). MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. *Oncogene* 36, 3346–3356.
- Pullarkat, V.A., and Newman, E.M. (2016). BCL2 Inhibition by Venetoclax: Targeting the Achilles' Heel of the Acute Myeloid Leukemia Stem Cell? *Cancer Discov* 6, 1082–1083.
- Rabbani, A., Finn, R.M., and Ausió, J. (2005). The anthracycline antibiotics: antitumor drugs that alter chromatin structure. *BioEssays* 27, 50–56.
- Rampal, R., Alkalín, A., Madzo, J., Vasanthakumar, A., Pronier, E., Patel, J., Li, Y., Ahn, J., Abdel-Wahab, O., Shih, A., et al. (2014). DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Reports* 9, 1841–1855.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8, 2281–2308.

Rao, R.C., and Dou, Y. (2015). Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nature Reviews Cancer* 15, 334–346.

Rapin, N., Bagger, F.O., Jendholm, J., Mora-Jensen, H., Krogh, A., Kohlmann, A., Thiede, C., Borregaard, N., Bullinger, L., Winther, O., et al. (2014). Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. *Blood* 123, 894–904.

Rath, D., Amlinger, L., Rath, A., and Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie* 117, 119–128.

Roe, J.-S., and Vakoc, C.R. (2016). The essential transcriptional function of BRD4 in acute myeloid leukemia cells. *Cold Spring Harb Symp Quant Biol* 81, 61–66.

Roe, J.-S., Mercan, F., Rivera, K., Pappin, D.J., and Vakoc, C.R. (2015). BET bromodomain inhibition suppresses the functional output of hematopoietic transcription factors in acute myeloid leukemia. *Mol Cell* 58, 1028–1039.

Rowley, J.D. (2008). Chromosomal translocations: revisited yet again. *Blood* 112, 2183–2189.

Saksouk, N., Avvakumov, N., Champagne, K.S., Hung, T., Doyon, Y., Cayrou, C., Paquet, E., Ullah, M., Landry, A.-J., Côté, V., et al. (2009). HBO1 HAT Complexes Target Chromatin throughout Gene Coding Regions via Multiple PHD Finger Interactions with Histone H3 Tail. *Molecular Cell* 33, 257–265.

Sánchez-Rivera, F.J., and Jacks, T. (2015). Applications of the CRISPR-Cas9 system in cancer biology. *Nature Reviews. Cancer* 15, 387–395.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* 32, 347–355.

Sauer, T., Arteaga, M.F., Isken, F., Rohde, C., Hebestreit, K., Mikesch, J.-H., Stelljes, M., Cui, C., Zhou, F., Göllner, S., et al. (2015). MYST2 acetyltransferase expression and Histone H4 Lysine acetylation are suppressed in AML. *Experimental Hematology* 43, 794-802.e4.

Saultz, J.N., and Garzon, R. (2016). Acute Myeloid Leukemia: A Concise Review. *Journal of Clinical Medicine* 5, 33.

Shalem, O., Sanjana, N.E., and Zhang, F. (2015). High-throughput functional genomics using CRISPR-Cas9. *Nature Reviews. Genetics* 16, 299–311.

Sharma, S., and Petsalaki, E. (2018). Application of CRISPR-Cas9 Based Genome-Wide Screening Approaches to Study Cellular Signalling Mechanisms. *Int J Mol Sci* 19.

Sharma, M., Zarnegar, M., Li, X., Lim, B., and Sun, Z. (2000). Androgen Receptor Interacts with a Novel MYST Protein, HBO1. *J. Biol. Chem.* 275, 35200–35208.

Sheikh, B.N., and Akhtar, A. (2018). The many lives of KATs — detectors, integrators and modulators of the cellular environment. *Nature Reviews Genetics* 1.

Shi, A., Murai, M.J., He, S., Lund, G., Hartley, T., Purohit, T., Reddy, G., Chruszcz, M., Grembecka, J., and Cierpicki, T. (2012). Structural insights into inhibition of the bivalent menin-MLL interaction by small molecules in leukemia. *Blood* 120, 4461–4469.

Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., and Vakoc, C.R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nature Biotechnology* 33, 661–667.

Shih, A.H., Jiang, Y., Meydan, C., Shank, K., Pandey, S., Barreyro, L., Antony-Debre, I., Viale, A., Socci, N., Sun, Y., et al. (2015). Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell* 27, 502–515.

Skucha, A., Ebner, J., Schmöllerl, J., Roth, M., Eder, T., César-Razquin, A., Stukalov, A., Vittori, S., Muhar, M., Lu, B., et al. (2018). MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity. *Nature Communications* 9, 1983.

Slany, R.K. (2009). The molecular biology of mixed lineage leukemia. *Haematologica* 94, 984–993.

Slany, R.K. (2016). The molecular mechanics of mixed lineage leukemia. *Oncogene* 35, 5215–5223.

Smith, E.R., Eisen, A., Gu, W., Sattah, M., Pannuti, A., Zhou, J., Cook, R.G., Lucchesi, J.C., and Allis, C.D. (1998). ESA1 is a histone acetyltransferase that is essential for growth in yeast. *PNAS* 95, 3561–3565.

Soutourina, J. (2018). Transcription regulation by the Mediator complex. *Nature Reviews Molecular Cell Biology* 19, 262–274.

Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., Vakoc, A.L., Kim, J.-E., Chen, J., Lazar, M.A., et al. (2008). DOT1L/KMT4 Recruitment and H3K79 Methylation Are Ubiquitously Coupled with Gene Transcription in Mammalian Cells. *Molecular and Cellular Biology* 28, 2825–2839.

Stein, E.M., and Tallman, M.S. (2015). Emerging therapeutic drugs for AML. *Blood* 127, 71–78.

Stein, E.M., Altman, J.K., Collins, R., DeAngelo, D.J., Fathi, A.T., Flinn, I., Frankel, A., Levine, R.L., Medeiros, B.C., Patel, M., et al. (2014). AG-221, an oral, selective, first-in-class, potent inhibitor of the IDH2 mutant metabolic enzyme, induces durable remissions in a Phase I study in patients with IDH2 mutation positive advanced hematologic malignancies. *Blood* 124, 115.

Stein, E.M., Garcia-Manero, G., Rizzieri, D.A., Tibes, R., Berdeja, J.G., Savona, M.R., Jongen-Lavrenic, M., Altman, J.K., Thomson, B., Blakemore, S.J., et al. (2018). The DOT1L inhibitor pinometostat reduces H3K79 methylation and has modest clinical activity in adult acute leukemia. *Blood* 131, 2661–2669.

Sugimoto, K., Toyoshima, H., Sakai, R., Miyagawa, K., Hagiwara, K., Ishikawa, F., Takaku, F., Yazaki, Y., and Hirai, H. (1992). Frequent mutations in the p53 gene in human myeloid leukemia cell lines. *Blood* 79, 2378–2383.

Tarumoto, Y., Lu, B., Somerville, T.D.D., Huang, Y.-H., Milazzo, J.P., Wu, X.S., Klingbeil, O., Demerdash, O.E., Shi, J., and Vakoc, C.R. (2018). LKB1, Salt-Inducible Kinases, and MEF2C Are Linked Dependencies in Acute Myeloid Leukemia. *Molecular Cell* 69, 1017–1027.e6.

TCGA Research Network (2013). Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N Engl J Med* 368, 2059–2074.

- Testa, J.R., Mintz, U., Rowley, J.D., Vardiman, J.W., and Golomb, H.M. (1979). Evolution of karyotypes in acute nonlymphocytic leukemia. *Cancer Research* 39, 3619–3627.
- The Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *New England Journal of Medicine* 368, 2059–2074.
- Thiel, A.T., Blessington, P., Zou, T., Feather, D., Wu, X., Yan, J., Zhang, H., Liu, Z., Ernst, P., Koretzky, G.A., et al. (2010). MLL-AF9-Induced Leukemogenesis Requires Coexpression of the Wild-Type Mll Allele. *Cancer Cell* 17, 148–159.
- Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development* 136, 3131–3141.
- Ting, P.Y., Parker, A.E., Lee, J.S., Trussell, C., Sharif, O., Luna, F., Federe, G., Barnes, S.W., Walker, J.R., Vance, J., et al. (2018). Guide Swap enables genome-scale pooled CRISPR–Cas9 screening in human primary cells. *Nature Methods* 15, 941.
- Tyner, J.W., Tognon, C.E., Bottomly, D., Wilmot, B., Kurtz, S.E., Savage, S.L., Long, N., Schultz, A.R., Traer, E., Abel, M., et al. (2018). Functional genomic landscape of acute myeloid leukaemia. *Nature* 562, 526.
- Tzelepis, K., Koike-Yusa, H., De Braekeleer, E., Li, Y., Metzakopian, E., Dovey, O.M., Mupo, A., Grinkevich, V., Li, M., Mazan, M., et al. (2016). A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Reports* 17, 1193–1205.
- Voss, A.K., and Thomas, T. (2009). MYST family histone acetyltransferases take center stage in stem cells and development. *BioEssays* 31, 1050–1061.
- Wagner, A.H., Coffman, A.C., Ainscough, B.J., Spies, N.C., Skidmore, Z.L., Campbell, K.M., Krysiak, K., Pan, D., McMichael, J.F., Eldred, J.M., et al. (2016). DGIdb 2.0: mining clinically relevant drug-gene interactions. *Nucleic Acids Research* 44, D1036–44.
- Wang, P., Lin, C., Smith, E.R., Guo, H., Sanderson, B.W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L.M., et al. (2009a). Global Analysis of H3K4 Methylation Defines MLL Family Member Targets and Points to a Role for MLL1-Mediated H3K4 Methylation in the Regulation of Transcriptional Initiation by RNA Polymerase II. *Molecular and Cellular Biology* 29, 6074–6085.
- Wang, T., Yu, H., Hughes, N.W., Liu, B., Kendirli, A., Klein, K., Chen, W.W., Lander, E.S., Sabatini, D.M., Mazan, M., et al. (2017). Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic Ras. *Cell* 168, 890–903.
- Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40, 897–903.
- Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009b). Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes. *Cell* 138, 1019–1031.
- Waskow, C., Rahmig, S., and Cosgun, N. (2014). Kit Deficiency Regulates Stable Human Hematopoietic Stem Cell Engraftment in Mice. *Blood* 124, 653–653.

- Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The origin and evolution of mutations in acute myeloid leukemia. *Cell* 150, 264–278.
- Welch, J.S., Petti, A.A., Miller, C.A., Fronick, C.C., O’Laughlin, M., Fulton, R.S., Wilson, R.K., Baty, J.D., Duncavage, E.J., Tandon, B., et al. (2016). TP53 and Decitabine in Acute Myeloid Leukemia and Myelodysplastic Syndromes.
- Wong, P.G., Glozak, M.A., Cao, T.V., Vaziri, C., Seto, E., and Alexandrow, M.G. (2010). Chromatin unfolding by Cdt1 regulates MCM loading via opposing functions of HBO1 and HDAC11-geminin. *Cell Cycle* 9, 4351–4363.
- Wood, K., Tellier, M., and Murphy, S. (2018). DOT1L and H3K79 Methylation in Transcription and Genomic Stability. *Biomolecules* 8.
- Wouters, B.J., and Delwel, R. (2016). Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. *Blood* 127, 42–52.
- Wu, J.Q., and Snyder, M. (2008). RNA polymerase II stalling: loading at the start prepares genes for a sprint. *Genome Biology* 9, 220.
- Yagi, H., Deguchi, K., Aono, A., Tani, Y., Kishimoto, T., and Komori, T. (1998). Growth Disturbance in Fetal Liver Hematopoiesis of Mll-Mutant Mice. *Blood* 92, 108–117.
- Yan, M.S., Turgeon, P.J., Man, H.S.J., Dubinsky, M.K., Ho, J.J.D., El-Rass, S., Wang, Y.-D., Wen, X.-Y., and Marsden, P.A. (2018). Histone acetyltransferase 7 (KAT7)-dependent intragenic histone acetylation regulates endothelial cell gene regulation. *J. Biol. Chem.* jbc.RA117.001383.
- Yan, Y., Harper, S., Speicher, D.W., and Marmorstein, R. (2002). The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. *Nature Structural & Molecular Biology* 9, 862–869.
- Yan Matthew S., and Marsden Philip A. (2015). Epigenetics in the Vascular Endothelium. *Arteriosclerosis, Thrombosis, and Vascular Biology* 35, 2297–2306.
- Yokoyama, A. (2015). Molecular mechanisms of MLL-associated leukemia. *Int J Hematol* 101, 352–361.
- Yokoyama, A. (2017). Transcriptional activation by MLL fusion proteins in leukemogenesis. *Experimental Hematology* 46, 21–30.
- Yokoyama, A., and Cleary, M.L. (2008). Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes. *Cancer Cell* 14, 36–46.
- Yokoyama, A., Somervaille, T.C.P., Smith, K.S., Rozenblatt-Rosen, O., Meyerson, M., and Cleary, M.L. (2005). The Menin Tumor Suppressor Protein Is an Essential Oncogenic Cofactor for MLL-Associated Leukemogenesis. *Cell* 123, 207–218.
- Yokoyama, A., Lin, M., Naresh, A., Kitabayashi, I., and Cleary, M.L. (2010). A Higher-Order Complex Containing AF4 and ENL Family Proteins with P-TEFb Facilitates Oncogenic and Physiologic MLL-Dependent Transcription. *Cancer Cell* 17, 198–212.
- Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A.J., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* 378, 505.

Yuan, H., Rossetto, D., Mellert, H., Dang, W., Srinivasan, M., Johnson, J., Hodawadekar, S., Ding, E.C., Speicher, K., Abshiru, N., et al. (2012). MYST protein acetyltransferase activity requires active site lysine autoacetylation. *The EMBO Journal* 31, 58–70.

Zeisig, B.B., and So, C.W.E. (2015). Cellular and Molecular Basis of KMT2A/MLL Leukaemias: From Transformation Mechanisms to Novel Therapeutic Strategies. In *Chromosomal Translocations and Genome Rearrangements in Cancer*, J.D. Rowley, M.M. Le Beau, and T.H. Rabbitts, eds. (Cham: Springer International Publishing), pp. 223–250.

Zhang, T., Cooper, S., and Brockdorff, N. (2015). The interplay of histone modifications – writers that read. *EMBO Reports* 16, 1467–1481.

Zhang, W., Prakash, C., Sum, C., Gong, Y., Li, Y., Kwok, J.J.T., Thiessen, N., Pettersson, S., Jones, S.J.M., Knapp, S., et al. (2012). Bromodomain-containing Protein 4 (BRD4) Regulates RNA Polymerase II Serine 2 Phosphorylation in Human CD4+ T Cells. *J. Biol. Chem.* 287, 43137–43155.

Zhang, Y., Liu, T., Meyer, C.A., Eickhout, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9, R137.

Zhu, L., Li, Q., Wong, S.H.K., Huang, M., Klein, B.J., Shen, J., Ikenouye, L., Onishi, M., Schneidawind, D., Buechele, C., et al. (2016). ASH1L Links Histone H3 Lysine 36 di-methylation to MLL Leukemia. *Cancer Discov* 6, 770–783.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., et al. (2011). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* 478, 524–528.